

FRANCISCO DE OLIVEIRA MENESES

**“AS VOGAIS DESVOZEADAS NO PORTUGUÊS BRASILEIRO: INVESTIGAÇÃO
ACÚSTICO-ARTICULATÓRIA”**

**“DEVOICED VOWEL IN BRAZILIAN PORTUGUESE: AN ACOUSTIC-
ARTICULATORY INVESTIGATION”**

**Campinas
2012**

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE ESTUDOS DA LINGUAGEM

Francisco de Oliveira Meneses

**AS VOGAIS DESVOZEADAS NO PORTUGUÊS BRASILEIRO: INVESTIGAÇÃO
ACÚSTICO-ARTICULATÓRIA**

Orientador: Eleonora Cavalcante Albano

***DEVOICED VOWELS IN BRAZILIAN PORTUGUESE: AN ACOUSTIC-
ARTICULATORY INVESTIGATION***

Adviser: Eleonora Cavalcante Albano

Dissertação de mestrado apresentada ao Instituto de Estudos da Linguagem da Universidade Estadual de Campinas para a obtenção do título de Mestre em Linguística.

Dissertation submitted to the Institute of Language Studies of the State University of Campinas for obtaining the Master's degree in Linguistics.

Campinas, 2012

FICHA CATALOGRÁFICA ELABORADA POR
TERESINHA DE JESUS JACINTHO – CRB8/6879 - BIBLIOTECA DO INSTITUTO DE
ESTUDOS DA LINGUAGEM - UNICAMP

M524v Meneses, Francisco, 1986-
As vogais desvozeadas no Português Brasileiro:
investigação acústico-articulatória / Francisco de Oliveira
Meneses. -- Campinas, SP : [s.n.], 2012.

Orientador : Eleonora Cavalcante Albano.
Dissertação (mestrado) - Universidade Estadual de
Campinas, Instituto de Estudos da Linguagem.

1. Vogais. 2. Fonologia acústico-articulatória. 3. Alofonia
(Fonologia). 4. Fonética. 5. Fonologia gestual. I. Albano,
Eleonora Cavalcante, 1950-. II. Universidade Estadual de
Campinas. Instituto de Estudos da Linguagem. III. Título.

Informações para Biblioteca Digital

Título em inglês: Devoiced vowels in Brazilian portuguese: an acoustic-articulatory investigation.

Palavras-chave em inglês:

Vowels

Acoustic-Articulatory Phonology

Alofonia (Fonologia)

Phonetics

Gestural phonology

Área de concentração: Linguística.

Titulação: Mestre em Linguística.

Banca examinadora:

Eleonora Cavalcante Albano [Orientador]

Vera Pacheco


Wilmar Rocha D'Angelis

Data da defesa: 04-06-2012.

Programa de Pós-Graduação: Linguística.

BANCA EXAMINADORA:

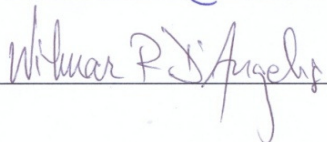
Eleonora Cavalcante Albano



Vera Pacheco



Wilmar Rocha D'Angelis



Regina Célia Fernandes Cruz

Maria Filomena Spatti Sândalo

IEL/UNICAMP
2012

Dedico este trabalho a meus amigos, Mateus e Pedro, que partiram cedo demais.

À minha esposa, principal razão para persistir e prosseguir.

AGRADECIMENTOS

À professora Eleonora Cavalcante Albano, que me recebeu em 2010 e foi mais do que orientadora nestes últimos anos. Foi responsável por meu crescimento acadêmico e pessoal. Agradeço imensamente por suas contribuições e por sua paciência.

À Prof^a. Vera Pacheco, que me introduziu no caminho da pesquisa, agradeço o carinho, os conselhos e o apoio irrestrito.

Aos sujeitos desta pesquisa, agradeço a simpatia e a disposição que demonstraram durante as sessões de gravação.

Aos membros da banca de qualificação, Prof. Dr. Wilmar Rocha D'Angelis e Prof^a. Dr^a. Regina Cruz, agradeço a leitura precisa e contribuições. Agradeço ainda a Prof^a. Dr^a. Filomena Sândalo, que aceitou prontamente o convite para a banca da defesa.

À Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP, pelo apoio financeiro.

Aos professores do Departamento de Linguística do IEL, que contribuíram para minha formação durante os últimos anos.

Aos funcionários do IEL, da Secretaria de Pós-Graduação pela atenção com que sempre nos atendem.

Agradeço aos professores e a todos os meus amigos e colegas da graduação em Letras Vernáculas da Universidade Estadual do Sudoeste da Bahia - UESB.

A todos os meus colegas do Mestrado em Linguística, por terem sido os meus primeiros companheiros na UNICAMP.

Aos colegas do LAFAPE: Luciana Lessa, Denise Pozzani, Maria Claudia de Freitas, Larissa Rinaldi, Antonio Pessoti, Laudino Roces, Lucila Schieleman, Diego Jiquilin Ramirez e Magnun Madruga que sempre tornaram o ambiente de estudo mais descontraído.

À D. Dulce Valle por ter me acolhido em Campinas durante a seleção do mestrado.

Agradeço aos meus amigos e colegas da vida: Deborah, Beatriz, Cirlene, Leandro, Audinéia, Milena Barbosa, Ivana Ivo, Daiane Rodrigues, Joselito Junior, Lázaro, Tia Celeste, Cátia, Samara, Jamile, Geová Junior, Ingrid Bahia, Gabriel Leite, Cássio, Juarez Alves. Muito obrigado pelas conversas que salvaram muitos dos meus dias.

Agradeço às minhas amigas Tatiane Macedo e Crislayne Alfagalli pela amizade; pelos momentos felizes de nossos cafés e pelos momentos em que sofremos as angústias da academia e as saudades de casa. Sem vocês tudo teria sido tudo muito sem graça!

Agradeço aos meus avós, tios, primos pelo apoio de toda vida. Em especial, agradeço ao meu primo e amigo Danilo com quem compartilhei alegrias e medos de minha empreitada. Agradeço também aos meus sogros, Elza e Augusto, e minha cunhadinha, Patrícia, pelo apoio dispensado em todos esses anos.

A meus pais, Norma e João, e a minha irmã, Ana Clara, por me apoiarem e incentivarem em todas as decisões que tomei até aqui. Essa conquista é fruto da dedicação e do amor de nossa família.

Por fim, meu agradecimento e amor incondicional a Juliana, minha esposa, amiga e companheira de todas as horas. Obrigado pelo incentivo, pela paciência nas noites mal dormidas, pela troca de ideias e pelo apoio em todo o percurso dedicado a este trabalho.

RESUMO

Este trabalho examina o processo de produção de vogais desvozeadas no português brasileiro (doravante PB). O objetivo é, a partir de uma visão de cunho dinâmico, mostrar que há, no sinal das chamadas sílabas desvozeadas, rastros da vogal que impossibilitam uma hipótese de síncope vocálica. Além disso, busca-se uma síntese teórica das pistas acústicas encontradas a fim de relacioná-las à articulação das vogais desvozeadas. Para tanto, uma análise acústico-articulatória foi realizada a partir da gravação da leitura de frases-veículos. As leituras das frases foram realizadas por seis sujeitos do sexo feminino, naturais de Vitória da Conquista – BA, sem queixas de fala ou audição. As gravações foram realizadas em uma cabine acústica, por meio de gravador digital. A análise instrumental foi realizada por meio do software PRAAT. Foram obtidas as medidas de duração da sílaba e do ruído fricativo, as medidas do centroide do ruído e os valores da Razão de centralização formântica (doravante FCR) e a Área de espaço vocálico (doravante VSA) das vogais em contexto de desvozeamento. Os resultados encontrados mostram que há um gradiente de desvozeamento manifestado de três maneiras: as medidas de duração, assim como as medidas do primeiro momento espectral, mostraram que pistas remanescentes do gesto vocálico permanecem no ruído das fricativas; os dados de FCR e VSA mostram que as vogais sofrem uma grande redução da magnitude em contexto de desvozeamento. Os resultados apontam para um efeito de sobreposição de gestos, em detrimento de uma síncope vocálica. Essa hipótese não caberia em descrições fonológicas tradicionais, as quais, em geral, lidam com operações simbólicas. O fenômeno em estudo pode, então, ser iluminado pela Fonologia Gestual (BROWMAN E GOLDSTEIN, 1992; BALL E KENT, 1997; ALBANO, 2001), a qual, em relação à representação tradicional, é capaz de expressar realizações gradientes, pois incorpora com sucesso os fatores tempo e magnitude, diretamente relacionados à ideia de movimento dos articuladores.

Palavras-Chave: Vogais desvozeadas, Fonética, Fonologia Gestual, Gradientes alofônicos.

ABSTRACT

This paper examines the production of devoiced vowels in Brazilian Portuguese. The goal is to depart from a dynamic view to show that there are vowel traces in the signal of devoiced syllables, a fact which weakens the hypothesis of vowel syncope. In addition, we seek a theoretical synthesis of the acoustic cues found in order to relate them to devoiced vowel articulation. To this end, an acoustic-articulatory analysis was performed from the recording of read carrier sentences. The sentences were read by six female subjects, from Vitória da Conquista - BA, without speech or hearing problems. The recordings were performed in an acoustic booth, using a digital recorder. The instrumental analysis was performed using the PRAAT software. The following measurements were made: the duration of the syllable and its fricative noise, the centroid of the noise and the values of formant centralization ratio (FCR) and vowel space area (VSA) of vowels in devoicing context. The results show that there is gradient devoicing manifested in three ways: both the duration measurements and the first spectral moment measures showed that the vowel gesture cues remain in fricative noise signal; FCR and VSA data show that vowels undergo great magnitude reduction in the devoicing context. The results indicate an effect of overlapping gestures, rather than vowel syncope. This hypothesis would not fit into traditional phonological descriptions, which usually deal with symbolic operations. The phenomenon under study can thus then be illuminated by Gestural Phonology (BROWMAN AND GOLDSTEIN, 1992; BALL AND KENT, 1997; ALBANO, 2001), which, as opposed to traditional representation, is able to express gradients, as it incorporates successfully the factors of timing and magnitude, directly related to the idea of articulator movement.

Key-Words: Devoiced Vowels, Phonetics, Gestural Phonology, Allophonic gradients.

LISTA DE TABELAS

Tabela 1:	Frequência de ocorrência das palavras alvo no PB.	47
Tabela 2:	Análise de variância da duração das sílabas /s/ + vogal /u/ vozeada, /s/ + vogal /u/ desvozeada parcial e /s/ + vogal /u/ desvozeada total.	53
Tabela 3:	Análise de variância da duração das sílabas /s/ + vogal /i/ vozeada, /s/ + vogal /i/ desvozeada parcial e /s/ + vogal /i/ desvozeada total.	54
Tabela 4:	Análise de variância da duração do /s/ que acompanha vogal /i/ vozeada, parcialmente desvozeada e totalmente desvozeada	55
Tabela 5:	Análise de variância da duração do /s/ que acompanha vogal /u/ vozeada, parcialmente desvozeada e totalmente desvozeada.	57
Tabela 6:	T pareado para valores do centróide de /i/ desvozeado e /i/ vozeado	60
Tabela 7:	Teste T pareado para valores do centróide de /i/ desvozeado e /i/ vozeado.	61
Tabela 8:	Valores de R e p para correlação entre medidas de FCR e VSA das vogais em contexto de desvozeamento	68

LISTA DE FIGURAS

Figura 1:	Duração de /s/ e /l/ com schwa vozeado vs. schwa desvozeado em três condições de velocidade de fala retirado de Davidson	9
Figura 2:	Representação da atividade glotal retirada de Yoshioka (1981)	11
Figura 3:	Frequência de ocorrência de desvozeamento por vogal no experimento piloto	29
Figura 4a:	Rótulos em <i>Text Grid</i> para dados referentes à sequência fricativa + vogal desvozeada/vozeada.	31
Figura 4b:	Rótulos em <i>Text Grid</i> dos dados referente à sequência fricativa + vogal vozeada.	32
Figura 5:	Ausência de periodicidade e de limites entre fricativa e vogal no sinal acústico.	33
Figura 6:	Janela de objetos do Praat para a obtenção do valor referente ao centróide, desvio padrão, assimetria e curtose.	36
Figura 7:	Exemplo de /i/ vozeado em (a) e /i/ desvozeado em (b) para a palavra “lance” /ˈlaNsi/. As porções destacadas indicam em (a) periodicidade da vogal vozeada e em (b) a extensão do desvozeamento total.	41
Figura 8:	Exemplo de /a/ parcialmente desvozeado para a palavra “caça” /ˈkasa/.	42
Figura 9:	Ocorrências de desvozeamento total (DV total), desvozeamento parcial (DV Parcial) e de vogais vozeadas em todo conjunto de dados.	43
Figura 10:	Taxa de desvozeamento total (DV total), desvozeamento parcial (DV Parcial) e vogal vozeada para /u/, /a/ e /i/.	44
Figura 11:	Taxa de realização de desvozeamento total (DV total), desvozeamento parcial (DVParcial) e vogais vozeadas para	45

cada sujeito analisado.

Figura 12:	Percentagem de vogais desvozeadas por sujeito.	46
Figura 13:	Taxa de desvozeamento total (DV total), desvozeamento parcial (DV Parcial) e vogais vozeadas para cada palavra.	46
Figura 14:	Duração das sílabas /s/ + vogal vozeada (at_s_u Voz), /s/ + vogais desvozeada parcial (at_s_u VDP) e /s/ + vogal desvozeada total (at_s_u VD) para a vogal /u/.	52
Figura 15 :	Duração das sílabas /s/ + vogal vozeada (at_s_i Voz), /s/ + vogais desvozeada parcial (at_s_i VDP) e /s/ + vogal desvozeada total (at_s_i VD) para a vogal /i/.	53
Figura 16:	Duração do /s/ que acompanha vogal /i/ vozeada (S_i_Voz), parcialmente desvozeada (S_i_VDP) e totalmente desvozeada (S_i_VD).	55
Figura 17:	Duração do /s/ que acompanha vogal /u/ vozeada (S_u_Voz), parcialmente desvozeada (S_u_VDP) e totalmente desvozeada (S_u_VD).	56
Figura 18:	Centróide de /s/ que acompanha a vogal /i/ desvozeada e /i/ vozeada com filtro de 500 Hz.	60
Figura 19:	Centróide de /s/ que acompanha a vogal /i/ desvozeada e /i/ vozeado.	61
Figura 20:	Espaço vocálico de vogais vozeadas em contexto de desvozeamento para sujeitos do PB	64
Figura 21:	Espaço vocálico de vogais vozeadas em contexto normal para sujeitos do PB.	64
Figura 22:	Valores médios da Área de espaço vocálico em contexto de desvozeamento e em contexto de vozeamento.	65
Figura 23:	Valores médios da Razão de Centralização Formântica em contexto de desvozeamento e em contexto de vozeamento.	67

Figura 24:	Diagrama de dispersão - FCR X VSA	68
Figura 25:	Proposta de redução vocálica em contexto de desvozeamento vocálico.	71
Figura 26:	Espectrograma de [si] desvozeado com banda de frequência de 0-8 KHz.	73

SUMÁRIO

INTRODUÇÃO.....	1
CAPÍTULO 1.....	5
1.1 O problema	5
1.1.1 Os dados do japonês.....	6
1.1.2 A diferença dos dados do francês.....	7
1.1.3 Os dados do inglês.....	8
1.2 O desvozeamento e as descrições do procedimento articulatório.....	10
CAPÍTULO 2.....	15
2.1 Sobreposição de gestos x apagamento vocálico.....	15
2.2 A fonologia gestual.....	20
CAPÍTULO 3 – Metodologia.....	25
3.1 Sujeitos e Coleta dos dados	25
3.2 Preparação do corpus.....	26
3.2.1 O estudo piloto: resumo dos resultados.....	27
3.2.2 A montagem do corpus final.....	29
3.3 Segmentação e rotulação dos dados.....	31
3.4 Análise dos dados	32
3.4.1 Medidas de duração.....	33
3.4.2 Medidas da Razão de Centralização Formântica e Área de Espaço Vocálico.....	34
3.4.3 Análise da forma do espectro.....	35
3.5 Análises Estatística.....	36
3.6 O uso de Scripts	37
CAPÍTULO 4 – Resultados e discussão I.....	39
4.1 Considerações Preliminares.....	39

4.2 Características das vogais desvozeadas do PB.....	40
4.3 Discussão	48
 CAPÍTULO 5 – Resultados e discussão II.....	 51
5.1 Queda ou sobreposição de gestos: a análise da duração.....	51
5.2 Pistas acústicas: a análise do centróide.....	58
5.3 VSA e FCR e a redução vocálica.....	62
5.4 Discussão	69
 CONSIDERAÇÕES FINAIS.....	 77
REFERÊNCIAS BIBLIOGRÁFICAS	81
ANEXOS	89

Introdução

*“Todo o conhecimento humano começou com intuições,
passou daí aos conceitos e terminou com ideias.”*

Immanuel Kant

O desvozeamento de vogais¹ é um processo recorrente em muitas línguas naturais. O japonês, o francês e o inglês, entre outras, são relatadas na literatura como línguas com vogais desvozeadas. O desvozeamento é usualmente observado quando as vogais altas [i] e [u], além do [ə] para o inglês, ocorrem entre segmentos surdos. Por exemplo, nas seguintes palavras do português brasileiro (doravante PB), as vogais sublinhadas podem ser desvozeadas: *timaço*, *basquete*, *aço* e *lance*.

Em geral, as vogais altas, que, sendo intrinsecamente breves, tendem a perder o vozeamento, o que se manifesta articulatoriamente como uma falta de vibração das pregas vocais e acusticamente como uma ausência de periodicidade no sinal acústico.

A literatura tem mostrado que as vogais surdas nem sempre têm todas as características de ressonâncias do trato vocal visíveis no espectrograma (HAN, 1962;

¹ Usamos já aqui o termo desvozeamento seguindo as descrições tradicionais desse processo. A natureza exata desse fenômeno no PB (desvozeamento ou síncope), no entanto, ainda será tratada e é objeto da presente dissertação de mestrado.

KAWAKAMI, 1977; BECKMAN, SHOJI, 1984). Assim, são percebidas como um longo intervalo de aspiração ou fricção após a liberação da consoante, resultando em limites indistintos entre vogais surdas e consoantes surdas adjacentes.

Beckman e Shoji (1984), assim como Jun et al (1998), afirmam que efeitos coarticulatórios no sinal da fricativa precedente fornecem pistas da presença da vogal. Além disso, Soli (1981), em estudo voltado para o inglês, mostra que há claras pistas da existência da vogal em contexto de vogais desvozeadas, tais como *art̩icle*, *ident̩ity*, eliminando-se, assim, a hipótese de síncope vocálica. Chitoran e Iskarous (2008) argumentam que o processo de desvozeamento pode ser considerado como um processo de sobreposição temporal de gestos articulatórios. Segundo os autores, o gesto da vogal não é excluído, mas sobreposto e conseqüentemente desvozeado, por exemplo, pelo gesto de uma consoante [s] precedente.

Para o PB, Lemle (1960) já alertava para que as vogais pudessem ser produzidas como variantes posicionais surdas, propriamente ditas, em certos contextos e registros. Além disso, Battisti e Hermans (2007), em trabalho sobre a palatalização em dialetos do Sul do Brasil, afirmam que em palavras tais como [tomaʈʃi], a vogal alta é surda, criando a impressão de uma africada longa. No entanto, não há nenhum estudo exclusivo e sistemático sobre o desvozeamento de vogais e os efeitos articulatórios decorrentes do fenômeno no PB.

Sabe-se que a produção de vogais desvozeadas é um fenômeno comum nas línguas e parece também ser no PB. Também é notório que o fenômeno vocálico em questão parece não ser tão categórico quanto as análises mais tradicionais têm afirmado.

Assim, o objetivo do presente estudo é investigar as vogais desvozeadas do Português Brasileiro. Mais especificamente, será discutida neste trabalho a natureza gradiente e a ocorrência desse tipo de vogal, com o objetivo de avaliar se essas vogais deixam de existir (síncope vocálica) ou são encobertas pelas consoantes adjacentes. Para isso, uma análise acústico-articulatória será realizada.

Para alcançar os objetivos aqui propostos, a dissertação está dividida em quatro capítulos.

No primeiro e segundo capítulo, as análises linguísticas e as questões em torno da investigação das vogais desvozeadas serão o foco: apresentamos um panorama dos estudos do desvozeamento vocálico em outras línguas e a abordagem desse fenômeno. Além disso, o capítulo traz à discussão aspectos relacionados especialmente ao fenômeno estudado, destacando-se também a importância da atividade articulatória e abordagens dinâmicas das questões da produção da fala.

O terceiro capítulo aborda o percurso metodológico feito durante a pesquisa. Lá, detalhamos a seleção dos sujeitos, as estratégias de gravação e as formas de análise dos dados.

No quarto capítulo apresentamos a descrição dos dados, e discutimos achados novos propiciados por nossa análise inferencial e as reflexões sobre eles.

Por fim, no capítulo final, discutem-se as implicações dos resultados obtidos e apresentam-se as conclusões e considerações relevantes para uma abordagem dinâmica dos estudos da fala e da linguagem.

Capítulo 1

Revisão de Literatura I: características gerais das vogais desvozeadas

1.1 O problema

O fenômeno do desvozeamento vocálico tem sido descrito desde a década de 1960 na literatura linguística e fonética, principalmente em estudos ligados ao japonês (HAN, 1962; GREENBERG, 1969; HIROSE, 1971; SUGITO, 1982; MAEKAWA, 1990;). Recentemente um significativo número de trabalhos tem explorado a existência ou não de vogais desvozeadas em diferentes contextos e nas mais diversas línguas naturais como, por exemplo, no turco (JANNEDY, 1995), no coreano (JUN, BECKMAN, 1993; MO, 2007), no francês (FAGYAL, MOISSET, 1999; SMITH, 2003) e no inglês (RODGERS, 1996; DAVIDSON, 2005), entre outras.

Essas análises, em conjunto, comprovam a recorrência de vogais desvozeadas sob diferentes condições (ANDREEVA, KOREMAN, 2008; KONDO, 1997; TSUCHIDA

1997, JUN, BECKMAN, 1993). Muitas vezes classificado como uma síncope (apagamento total de um gesto vocálico), o desvozeamento se mostra como um fenômeno complexo e recorrente nas línguas naturais.

Gordon (1998) apresenta informações sobre os padrões de vogais desvozeadas nas mais variadas línguas. Na grande maioria dos casos, o desvozeamento vocálico é descrito como um fenômeno gradiente, variável e não contrastivo que afeta principalmente vogais altas. As vogais altas e o schwa, cuja duração intrínseca tende a ser breve, muitas vezes se encontram desvozeadas ou apagadas em determinados ambientes consonantais (LEHISTE, 1970).

Dessa forma, a relação entre duração e ensurdecimento de vogais parece existir, uma vez que em línguas nas quais a duração da vogal é contrastiva, somente a vogal alta breve sofre desvozeamento, enquanto, em línguas acentuais, vogais altas só o sofrem em sílabas átonas (RODGERS, 1996; SMITH, 2003).

1.1.1 Os dados do japonês

Shiraishi (2003), em estudo restrito ao japonês, afirma que o desvozeamento de vogais médias e baixas é extremamente raro, restrito a poucas palavras, assim como o de vogais acompanhadas de consoantes vozeadas (YOSHIDA, SAGISAKA, 1990).

Para Hasegawa (1999), vogais altas podem ser completamente desvozeadas, sem alterar os significados e a percepção das palavras, e quando os falantes nativos são convidados a pronunciar cada palavra com muito cuidado, as vogais são quase sempre vozeadas. Além disso, o grau de ensurdecimento da vogal varia de indivíduo para indivíduo.

Ao que parece, a realização ou não de vogais desvozeadas está sujeita a efeitos da variação diatópica. Fujimoto (2004) afirma que as taxas de realização de vogais surdas entre os falantes de Osaka são menores do que entre os falantes do japonês de Tokyo. Segundo o autor, a análise da duração das vogais mostrou que as

vogais realizadas pelos falantes de Tokyo foram significativamente menores do que entre os falantes de Osaka. Esse dado, por outro lado, corrobora também a existência de uma íntima relação entre o desvozeamento de vogais e a sua duração. Conforme Fujimoto (2004), a duração de /i / na mora /ki / para os falantes de Tóquio era constantemente mais breve. Esta tendência não foi achada na mora /ke/, por exemplo. Isso, segundo o autor, indica, categoricamente, que falantes de Tóquio diferenciam a realização de vogais altas e não altas, de modo que a duração segmental em sílabas com vogais altas é controlada de modo especial.

Ainda sobre o japonês, em relação aos fatores prosódicos que condicionam o desvozeamento vocálico, Tsuchida (1997) mostra que o desvozeamento é inibido em fronteiras de palavras. Tal resultado indica que a posição da vogal na palavra ou na frase é fator relevante no processo de desvozeamento. Além disso, a posição final de palavras e a compressão decorrente da alta velocidade possibilita a sobreposição de gestos que provocam o desvozeamento (GORDON, 1998; SUGITO, 1982; MAEKAWA, 1990; SATO, 1989).

1.1.2 A diferença dos dados do francês

Para Smith (2003), a associação de desvozeamento com a proeminência distingue francês parisiense de outras línguas e dialetos que têm a vogal desvozeada. Em muitos outros idiomas, ela ocorre em ambientes prosodicamente fracos. Por exemplo, em japonês, as vogais mais prováveis de serem surdas são vogais altas átonas cercadas por consoantes surdas (VANCE, 1987). O desvozeamento no grego Moderno ocorre em vogais átonas (DAUER 1980), e em Português Europeu é também limitado a vogais altas, e para posições pré-pausa e não acentuadas (LAVER, 1994; CRUZ, FERREIRA, 1999). No dialeto franco-canadense de Quebec, vogais altas podem ser desvozeadas em meio de palavras quando são átonas (WALKER, 1984; DECHAINED, 1991), especialmente se rodeadas, também, por segmentos surdos (OUELLET *et al*, 1999).

Em todas estas línguas, conforme Smith (2003), o desvozeamento vocálico é uma forma de redução de vogal: no francês do Quebec, grego e português Europeu, as vogais podem ser totalmente apagadas (sic) e não apenas desvozeadas, e em japonês, as vogais desvozeadas são muitas vezes quase não realizadas acusticamente. Para a autora, o ensurdecimento é um processo opcional nessas línguas, na maioria das vezes com aplicação a vogais altas sem proeminência (acento ou ênfase).

O francês parisiense compartilha com outras línguas o fato de que as vogais altas são mais susceptíveis ao desvozeamento, mas parece difícil interpretar o ensurdecimento da vogal em francês como uma forma de redução de vogal, como nas outras línguas, já que aí o desvozeamento ocorre em uma posição de proeminência frasal. Segundo Smith (2003), isto sugere uma visão de ensurdecimento no francês falado em Paris bem diferente da interpretação geral de ensurdecimento em outros idiomas, que são intimamente relacionados ao acento.

1.1.3 Os dados do inglês

Para o inglês, as características das vogais seguem o que é comum na literatura: a segunda vogal das palavras *carpeting*, *multiple*, *blanketing* provavelmente é desvozeada (RODGERS, 1996; 1997). Segundo Rodgers (1996, 1997), tudo fica evidente no sinal acústico: a vogal “surda” não tem forma de onda definida, não mostra periodicidade e no espectrograma há uma explosão seguida por uma longa fase de fricção.

Davidson (2005) examinou um grande número de sequências do tipo / #C₁əC₂-/, a fim de explorar a natureza da queda do schwa em Inglês. Segundo a autora, as análises fonológicas que defendiam a elisão do schwa eram baseadas nas intuições dos pesquisadores e os estudos fonéticos que propunham que a elisão é caracterizada como sobreposição de gestos foram restritos a poucos registros de fala.

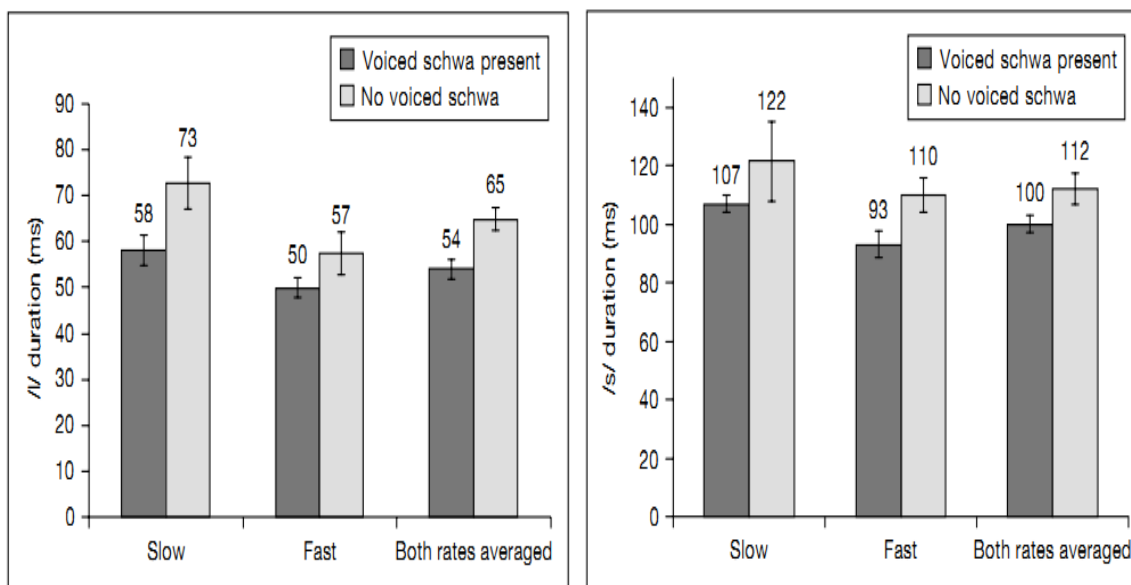


Figura 1: Duração de /s/ e /l/ com schwa vozeado vs. schwa desvozeado em três condições de velocidade de fala retirado de Davidson (2005, p. 91 e 92).

A duração do / l / é significativamente maior quando o schwa é elidido, em todas as taxa de elocução. Este achado, como pode ser visto na figura acima, é semelhante aos resultados para / s /, e consistente com a possibilidade de que o schwa seja, de fato produzido, mas encoberto no registro acústico.

Davidson (2005) afirma, com base nos dados acima, que a suposta “queda” do schwa é mais consistente com a hipótese de sobreposição gestual. Segundo ela, a produção do gesto de abertura da laringe pode encobrir ou mesmo impedir a produção do schwa vozeado².

Beckman (1996) tenta fornecer uma análise desses tipos de achados em termos de coordenação gestual. Após análises semelhantes em Browman e Goldstein (1990), Beckman (1996) argumenta que a aparente supressão do schwa em vários

² Tanto a questão da adoção dos gestos como unidade fonológica, quanto a hipótese do gesto de abertura da laringe encobrir o vozeamento das vogais serão discutidas em diferentes tópicos de nosso trabalho. Em relação ao último, não nos parece claro que há sobreposição na glote. Essa questão será discutida nos resultados e discussões, capítulo 4 da dissertação.

ambientes em Inglês poderia ser o desfecho extremo de um *continuum* de redução e sobreposição da vogal. Se assim for, ela observa que "os valores continuamente variáveis de sobreposição na pauta gestual são uma melhor representação do que a regra fonológica categórica de queda do schwa (p. 100). Além da queda do schwa no Inglês, Beckman (1996) também afirma que os fenômenos são similares no francês canadense, japonês, coreano.

1.2 O desvozeamento e as descrições do procedimento articulatório

Sabe-se que, durante a produção de um segmento vozeado, as pregas vocais vibram de maneira intermitente, isto é, aduzindo-se e abduzindo-se. Ao abrir-se, a glote permite que uma determinada quantidade de moléculas de ar passe, mas retém uma parte dessas moléculas ao fechar-se, imediatamente. Obviamente, quanto maior o número de moléculas de ar, mais elas se chocam, e maior é a quantidade de energia.

Em resumo, o vozeamento é produzido pela vibração das pregas vocais. Os requisitos principais para atingir o vozeamento na produção das vogais são: (i) as pregas vocais devem ser suficientemente aproximadas para que elas possam vibrar e (ii) a pressão subglótica deve exceder a pressão supraglótica por uma margem suficientemente grande. Se qualquer um desses requisitos deixar de ser atendido, o som produzido não será foneticamente vozeado.

As observações das ações glóticas para vogais produzidas no dialeto de Tokyo mostraram que, em vogais desvozeadas, a glote apresenta abertura contínua, sem os ciclos de fechamento. Em sequências CVC desvozeadas, a glote apresenta uma única ativação. As observações de Hirose (1971) revelam que a glote permanece aberta sem nenhum gesto de adução durante todo o período de produção da vogal desvozeada. Yoshioka (1981) também observa que a glote permanece aberta durante

toda a produção de vogais surdas, como pode ser visto na figura abaixo, e a atividade muscular é variável durante a produção de vogais desvozeadas.

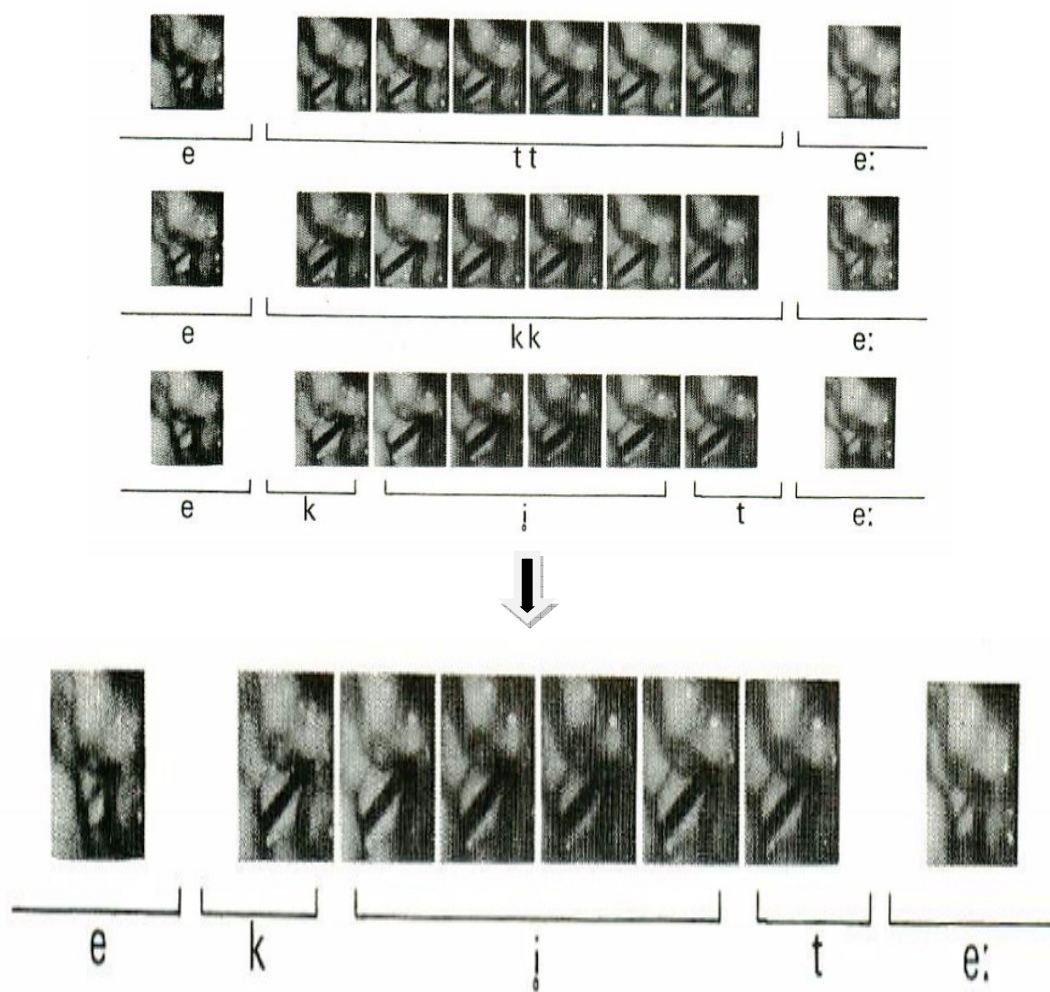


Figura 2: Representação da atividade glotal retirada de Yoshioka (1981).

A abertura da glote para [kit] e [kik] foi significativamente maior do que para outras sequencias testadas pelo autor. Segundo ele, isso mostra que os ajustamentos glotais para desvozeamento da vogal não são meros lapsos fonatórios na produção da vogal, mas um esforço positivo do alargamento da glote para segmentos surdos, ainda que não exista distinção fonêmica entre as vogais sonoras e surdas. Os dados,

como mostra a abertura glotal, também sugerem que a largura da glote para a vogal desvozeada também é influenciada consideravelmente pelas consoantes adjacentes.

Como pode ser visto através das figuras retiradas de Yoshioka (1981), a glote permanece aberta durante toda a produção de vogais desvozeadas, no caso o [i]. Segundo o próprio Yoshioka (1981), a atividade muscular é variável durante a produção destas vogais, o que sugere que o gesto é de alguma forma inibido, mas não completamente suprimido.

Para Ohala (1975), o desvozeamento significativamente maior em vogais altas do que em vogais baixas tem explicação fisiológica. Em conformidade com a previsão Ohala(1975), há uma maior queda de pressão através da glote em vogais abertas do que nas vogais fechadas. Outra consideração é que a elevação da língua que se encontra nas vogais altas como / i/ tende a tornar as pregas vocais mais tensas, por isso uma maior pressão subglótica é necessária para as pregas vocais começarem a vibrar (HONDA, 1983).

A explicação que Fagyal e Moisset (1999) propõem para o desvozeamento em seu experimento para o francês envolve uma violação do segundo dos requisitos que citamos acima. Eles sugerem que a constrição oral diminui durante as vogais surdas. O estreitamento da constrição oral permite que menos ar escape pela boca, assim, como resultado, mais ar torna-se preso na cavidade oral, aumentando a pressão supraglótica ao ponto de a diferença entre pressão subglótica e supraglótica não ser mantida.

Para Delforge (2008), como vogais altas são mais breves do que as vogais baixas ou médias em quase todas as línguas (LEHISTE, 1970), há uma maior probabilidade de que as abduções glóticas de consoantes surdas adjacentes impeçam a plena realização das aduções glóticas necessárias para a sonoridade vocálica. Além disso, a constrição oral estreita associada com a produção de vogais altas aumenta a pressão de ar na cavidade oral que inibe o fluxo de ar transglótico e, por conseguinte, torna as vogais mais susceptíveis ao ensurdecimento do que aqueles articulados com posições mais baixas da língua (JAEGER, 1978).

Rodgers (1996), em relação à proporção significativamente maior de desvozeadas em sílabas átonas do que em tônicas, afirma que sílabas tônicas são produzidas com maior força de articulação do que átonas: quanto maior a pressão subglótica em sílabas tônicas, a queda de pressão será suficiente em toda a glote para garantir vozeamento completo, o que não acontece, obviamente, nas átonas.

Capítulo 2

Revisão de Literatura II: entre o estático e o dinâmico

2.1 Sobreposição de gestos x apagamento vocálico

O título deste tópico, logicamente, antecipa nosso objetivo de discutir a relação apagamento/desvozeamento vocálico. Há diferenças óbvias e relevantes entre os termos: desvozeamento vocálico (DV) é o termo que aqui usamos para descrever o fenômeno sob investigação. Refere-se à falta de vozeamento da vogal e todos os atributos vocálicos visíveis que já mencionamos anteriormente. O apagamento vocálico, por outro lado, é o termo utilizado por investigadores que defendem a transcrição canônica, isto é, estática e que desconsidera as gradiências decorrentes da produção da fala. O termo designa “a ausência de uma secção vocálica segmentável no sinal de fala, ou seja, uma abertura oral associada com vozeamento, manifestando como uma estrutura típica” (HELGASON, KOHLER, 1996, p.143).

Por conta da divergência que tratamos acima, é comum na literatura que o processo que neste estudo assumimos como “desvozeamento vocálico” seja interpretado de duas maneiras: em uma, a vogal é analisada como excluída, como

Ohso (1973); na outra, a vogal é tratada como desvozeada em Han (1962) e Tsuchida (1997).

É importante lembrar que, quando vogais altas sofrem desvozeamento, faltam, no sinal acústico da vogal, componentes que são esperados para vogal tipicamente vozeada. Uma forma de onda periódica e regular e a barra de vozeamento são elementos que faltam às vogais surdas (BECKMAN, SHOJI 1984:63, HAN 1962, KONDO 2005, TSUCHIDA 1997). Além disso, a estrutura formântica em vogais que sofrem DV é irregular, como alguns investigadores observam (por exemplo: HAN, 1962). De acordo com Kondo (2005), a realização acústica de DV é como qualquer continuação de uma fricativa quando a vogal é precedida por uma fricativa.

Kawakami (1977), por exemplo, afirma que, a posição da língua ao término de [k] fornece pistas da presença das vogais [u] e [i] respectivamente. Segundo o autor, esta observação sugere que um rastro da vogal está presente na articulação, mas as pregas vocais não estão vibrando. No entanto, para Kawakami (1977), o mesmo não acontece para as fricativas.

Dadas as características variáveis elencadas acima, as análises de vogais que sofrem desvozeamento/apagamento são muitas vezes divergentes.

Beckman (1982, p. 118), examinando vogais e baseando-se na mora em japonês, inclui comparações da duração entre sílabas de CV cuja vogal sofre DV e sílabas de CV nas quais a vogal é sonora. A autora analisa a vogal em DV como apagamento. Beckman (1982) afirma que, ‘psicologicamente’, desvozeamento “está mais correto que apagamento porque os falantes do japonês ‘ouvem’ a vogal, mas apagamento está mais correto do que desvozeamento fisicamente”, pois geralmente não há nenhuma evidência espectral de existência da vogal. Por esse motivo, ela escolhe usar o “termo fisicamente mais preciso”, isto é, apagamento. Beckman e Shoji (1984) discutem a regra fonológica de apagamento vocálico de Ohso (1973), formulada abaixo:

V [+high] → Ø / [-voice] __ {[-voice], #}

Na análise de apagamento, a vogal perde suas características típicas quando é precedida por consoantes surdas. Os autores concluem que:

when the vowel is deleted for a devoiced syllable,[...], it may be only the time portion associated with the vowel as a laryngeal gesture that is removed. A portion of the lingual gesture, if it is compatible with the gesture toward the following consonant, may be maintained (Beckman, Shoji, 1984, p. 70).

Com respeito ao aspecto temporal, nos resultados de Beckman (1982), as medidas de duração típicas não apontam para a existência de uma porção vocálica seguindo a consoante.

Kondo (1997), embora não discuta o estado da vogal por si só, apresenta uma análise de queda vocálica. Ela analisa o desvozeamento de vogais como resultado de uma “desmoraficação”. Segundo a autora, o desvozeamento é seguido pela queda da sílaba. Essa queda é baseada na perda de sonoridade na sílaba: quando a vogal alta /i/ se torna surda, a mora não vozeada fica impedida de ser anexada ao nó da sílaba, pois a sílaba perdeu seu elemento central. O "elemento central" pode ser interpretado como o vozeamento.

Em suma, vários pesquisadores analisam a vogal desvozeada como uma apócope vocálica. Ohso (1973) trata a vogal como excluída com base na ausência acústica de vogal e através de uma análise tradicional; e Kondo (1997), como resultado da queda da mora.

Por outro lado, vários autores analisam a vogal surda não como uma síncope, mas, sim, como um desvozeamento. Por exemplo, McCawley (1968) propõe que as

vogais /i/ e /u/ tornam-se desvozeadas através de sua regra de ensurdecimento dada abaixo:

$$[-\text{cns}, +\text{voc}, +\text{dif}] \longrightarrow [-\text{voice}] / [-\text{voice}] __ \{[-\text{voice}], \#\}$$

Tsuchida (1997) examinou a sequência [sVc] onde V é uma vogal alta e C é um de [t, d, k, g] e encontrou nos sinais acústicos efeitos de coarticulação entre a vogal e a consoante seguinte. Quando a consoante era desvozeada a identificação da vogal diminui, mas as características acústicas permanecem.

Tsuchida (1997) argumenta a favor da análise de desvozeamento em detrimento de eliminação. A autora relata que no Japonês o /si/ com /i/ vozeado tem centróides maiores do que sílabas com /i/ desvozeado. Tsuchida (1997) afirma que a pista da vogal aparece apenas na última parte da fricativa nos dados acústicos.

Torreirar e Ernestus (2010) investigam o desvozeamento de vogais no francês europeu. Através de dados de fala espontânea e da análise do centróide das fricativas, os pesquisadores afirmam que o ensurdecimento é um fenômeno frequente nas vogais altas precedidas por consoantes mudas. Além disso, também mostram que o ensurdecimento é mais frequente nas vogais temporalmente reduzidas e coarticuladas.

Segundo os autores, no caso de /ti/ e /si/, o aumento da palatalização de /t/ e /s/ deve resultar em um deslocamento para baixo dos parâmetros do centróide. Para o /si/, os autores encontraram uma diferença significativa na direção predita pela hipótese.

Abordagens mais recentes tratam o ensurdecimento da vogal como resultado de processos articulatórios. Jannedy (1995), bem como Jun *et al* (1998), argumentam que o desvozeamento das vogais, em turco e coreano, respectivamente, é resultado de sobreposição gestual das vogais com suas consoantes desvozeadas adjacentes. Isto é, seu argumento é que o gesto da glote para as vogais altas é

sobreposto por conta dos gestos das consoantes surdas, resultando, assim, no ensurdecimento da vogal alta.

Jun e Beckman (1993) afirmam que alguns gestos das vogais altas do japonês também são sobrepostos. Segundo os autores, o gesto de fechamento glótico é encoberto pelos gestos de abertura da glote das consoantes surdas adjacentes. Os autores ainda afirmam que a sobreposição é mais comum entre vogais altas por conta da duração reduzida.

Em resumo, ao analisar o status desse tipo de vogal no japonês, McCawley (1968) e Tsuchida (1997) a tratam como desvozeada, não excluída. Jun e Beckman (1993) e Jannedy (1995) também interpretam a vogal como desvozeada, mas, além disso, também afirmam que o desvozeamento é um processo de sobreposição gestual.

Por fim, Chitoran e Iskarous (2008), em uma abordagem de cunho dinâmico, buscaram determinar se vogais altas pré-tônicas em Lezgi sofrem síncope ou se desvozeiam. Para os autores, o gesto vogal não é excluído, mas é sobreposto e conseqüentemente inibido pelo gesto precedente da consoante [s].

De acordo com os autores, três resultados sustentam esta hipótese e, conseqüentemente, a interpretação do desvozeamento, em detrimento da síncope: energia média mais baixa de [s] antes de [u], maior energia do [s] antes de [i] em contraste com [a], e maior de energia em [s] antes do [i] átono em contraste com [a] acentuado.

Como podemos perceber, o desvozeamento de vogais, além de ser muito variável, como já foi dito, tem sido muito discutido. No entanto, essa discussão nunca recebeu a devida atenção para o português brasileiro. Para o PB, muitas perguntas, ou melhor, quase todas as perguntas permanecem sem resposta, perguntas que abordo neste trabalho. Para respondê-las, entretanto, faz-se necessário o aporte de uma teoria que seja capaz de explicar processos fônicos da fala rápida e capaz de explicar as variações do fenômeno em questão. Essa teoria é o objeto de nosso próximo tópico.

2.2 A fonologia gestual

Vários estudos (BECKMAN, 1994; JANNEDY, 1995; JUN, BECKMAN, 1993) abordam o desvozeamento da vogal alta dentro do quadro teórico da Fonologia Articulatória (BROWMAN, GOLDSTEIN, 1989 e seguintes), que tem os gestos como as unidades básicas de análise.

A base para este trabalho se encontra na FAAR – Fonologia Acústico-Articulatória (ALBANO, 2001), proposta que tornou a análise instrumental uma ferramenta de acesso a informações articulatórias. Essa teoria é uma versão modificada da Fonologia Articulatória (doravante FAR), proposta por Browman & Goldstein (1992), que adota como unidade de análise o chamado gesto articulatório. Mais recentemente, esse tipo de modelo vem sendo denominado de Fonologia Gestual, doravante FonGest.

A FonGest parte de três hipóteses principais sobre a natureza das unidades fonológicas que permitem que essas unidades desempenhem o seu papel duplo como unidades de ação e unidades de combinação (e contraste). São eles:

(i) that vocal tract activity can be analyzed into constriction actions of distinct vocal organs, (ii) that actions are organized into temporally overlapping structures, and (iii) that constriction formation is appropriately modeled by dynamical systems. (GOLDSTEIN, FOWLER, 2003, p. 161)

Em modelos dinâmicos de produção da fala, como a FonGest, os gestos articulatórios surgem como unidades fonológicas e, ao mesmo tempo, unidades de ação articulatória (físicas). As unidades na Fonologia Gestual são unidades de ação e, como consequência, não estáticas, e potencialmente sobrepostas (BROWMAN, GOLDSTEIN, 1992). De acordo com Browman e Goldstein (1992) o objetivo de um gesto é produzir uma constrição no trato vocal utilizando de um número de órgãos.

Estes órgãos incluem: lábios, ponta / lâmina da língua, o corpo da língua, a raiz da língua, glote e palato. Um gesto, então, é definido como ação de constrição de um dos órgãos do trato vocal. No caso do gesto de fechamento de lábio, por exemplo, três articuladores estão envolvidos: o lábio superior, o lábio inferior, e a mandíbula.

Assim, cada ação ocorrida durante a produção de fala envolve ao menos um gesto, implementado por um órgão constritor. Gestos diferentes se sobrepõem no tempo integrando-se a movimentos subsequentes dentro do domínio de unidades maiores de produção de fala, tais como as palavras. Assim, a variabilidade observada na atividade do trato vocal é decorrente da sobreposição de gestos no tempo, além da variação de magnitude gestual (ALBANO, 2001).

Segundo Silva (2002), modelos dinâmicos, como a FonGest, têm grande sucesso em explicar processos fônicos da fala rápida, tais como enfraquecimentos e apagamentos de segmentos, justamente por serem capazes de mostrar a existência de gradientes fônicos que não são representados pelas análises tradicionais, uma vez que o gesto articulatório descreve tanto características espaciais como temporais da fala, oferecendo as mesmas vantagens das fonologias não-lineares e, além disso, simplifica a descrição de segmentos.

Em relação a teorias simbólicas estáticas como as propostas de Ohso (1973) e McCawley (1968), a análise dinâmica traz para a teoria fonológica gradientes e contextos diferentes de produção. Assim, embora os estudos baseados nos modelos fonológicos tradicionais tenham trazido muitas contribuições para a compreensão de vários aspectos intrigantes da teoria fonológica, acreditamos que uma categorização estática das estruturas linguísticas, tal como proposta por esses modelos, acarreta a desconsideração de processos gradientes relevantes, ou seja, características acústicas e/ou articulatórias sensíveis na produção e percepção de segmentos. A FonGest assume que as estruturas físicas e cognitivas, consideradas distintas em outros modelos, estão em um mesmo sistema.

A produção da fala na FonGest toma por base o modelo da dinâmica de tarefas, proposto por Saltzman e Kelso (1987). Esse modelo permite que se calculem as trajetórias dos articuladores no tempo, coordenando sua estrutura.

Nossa investigação adotou, então, o gesto articulatório como unidade, pois se trata de um primitivo que permite não só a incorporação do detalhe fonético na descrição, mas também permite a representação de informações gradientes e contínuas no sistema fonológico (BROWMAN, GOLDSTEIN, 1992).

Ao tratar da produção e percepção das ações vocais (GOLDSTEIN, FOWLER, 2003), o modelo consegue abarcar em uma mesma descrição alofônica processos categóricos e gradientes, isto é, aqueles que envolvem relações temporais de sobreposição e os ajustes articulatórios neles envolvidos.

A proposta do presente estudo é averiguar o processo de desvozeamento vocálico partindo de um modelo dinâmico de produção da fala. Em contraposição aos modelos fonológicos tradicionais, propõe a utilização de informação acústica para a análise de diversos fenômenos.

Em geral, afirma-se que para o japonês, coreano, inglês, francês e outras línguas, o desvozeamento da vogal ocorre como resultado do aumento das sobreposições gestuais entre gestos vocálicos e consonantais. No entanto ainda existem lacunas que os trabalhos baseados na FonGest não esgotaram, tais como: qual o percurso articulatório envolvido na produção de vogais desvozeadas? Quais gestos vocálicos são sobrepostos no desvozeamento?

Face ao exposto, procura-se explicar aqui como as vogais desvozeadas são produzidas, como são gradualmente implantadas como variantes e como se dá articulação dos gestos envolvidos na produção de tais vogais. Ainda, estudou-se em que medida os dados acústicos podem espelhar as estratégias articulatórias utilizadas pelos falantes do PB. A tentativa de nosso trabalho é, a partir de uma análise de cunho dinâmico, mostrar que há, no sinal das chamadas sílabas desvozeadas, rastros da vogal que impossibilitam uma hipótese de apócope vocálica. Além disso, busca-se

sintetizar as pistas acústicas encontradas aqui e relacioná-las à articulação das vogais desvozeadas e também relacioná-las aos achados de outras línguas.

Capítulo 3

Metodologia

3.1 Sujeitos e Coleta dos dados

Os sujeitos desta pesquisa foram escolhidos aleatoriamente durante o primeiro semestre de 2010. Por conta da facilidade, as informantes foram selecionadas no campus da Universidade Estadual do sudoeste da Bahia – UESB. Foram selecionadas seis informantes (JN, BR, BAR, DN, ANM, TC) naturais de Vitória da Conquista – Bahia. Para fins de padronização, todas as informantes são do sexo feminino. Além disso, todas são universitárias e têm a mesma faixa etária (20 - 30 anos). Os sujeitos não apresentam nenhum tipo de queixa fonoaudiológica.

O experimento foi realizado em duas sessões (3 informantes para cada bloco). As coletas de dados das três primeiras informantes ocorreram entre maio e junho de 2010, e a segunda coleta, de junho a julho de 2011. As gravações foram feitas em sessão única para cada falante, com pausas a critério dos sujeitos. As sessões duraram, em média, 40 min. cada. A ordem de classificação dos sujeitos (de 1 a 6) não seguiu a ordem de gravação, pois foi feita de forma aleatória.

As gravações ocorreram na cabine acústica do Laboratório de Fonética e Fonologia da Universidade Estadual do Sudoeste da Bahia – UESB, campus de Vitória da Conquista. A taxa de amostragem foi de 22 KHz, com 5 repetições de cada sentença. As fichas com as palavras foram embaralhadas antes de cada repetição com o objetivo de evitar a memorização, caso fosse mantida sempre a mesma ordem.

Os equipamentos utilizados foram: um computador com processador Intel Dual Core 2.4 GHZ, 4 GB DDRII 667 MHZ, HD 250 GB, 1 HD externo SAMSUNG GB (BACKUP), DVD-RW; 1 microfone headset Philips. O programa de gravação utilizado foi o Audacity (software livre). O programa *Praat* foi utilizado para fazer as medidas e as grades de texto (*text grids*) com as transcrições. Essas foram feitas manualmente em cada arquivo gerado.

3.2 Preparação do corpus

No período inicial de elaboração da dissertação de mestrado, foi proposto um grupo de palavras a ser posteriormente reavaliado. Esse primeiro corpus era composto por dissílabos, no quais as vogais /i,u,a/ ocupavam as posições de núcleo da última sílaba, por serem as únicas que ocorrem em posições finais no PB. Já as posições consonantais eram ocupadas por oclusivas (/t,d/) e fricativas (/s,z/). As palavras foram escolhidas a partir de corpus anteriormente preparado no LAFAPE seguindo o princípio das posições descritas acima.

Com o avanço das leituras e revisão da literatura, percebeu-se que as palavras pré-selecionadas seriam incapazes de contemplar e alcançar os objetivos propostos no projeto inicial, uma vez que a maioria delas não favorecia a ocorrência do desvozeamento vocálico conforme descrito na literatura. A maioria das palavras reais elencadas era composta por consoantes sonoras e não existia a possibilidade da contraparte surda. O desvozeamento total é, geralmente, observado quando as

vogais altas ocorrem entre segmentos desvozeados (CHITORAN, ISKAROUS, 2008; SMITH, 2010).

Com objetivo de averiguar que posições acentuais do PB favorecem, com maior ou menos intensidade, o desvozeamento de vogais foi montado um corpus piloto. Este corpus, com um número pequeno de palavras, era composto de trissílabos e tinha um total de 90 realizações. As sílabas C1+V contavam com uma consoante fricativa/africada seguida da vogal /i/ e ocupavam a posição tônica, pré-tônica e pós-tônica, como por exemplo, *chileno*, *cochilo*, *boliche*. Tais palavras foram aleatorizadas e inseridas numa frase veículo de fácil pronúncia. É importante ressaltar que, para esse primeiro experimento, optamos por realizar as gravações por meio de repetições das frases veículos e não apenas das palavras isoladas, para tentar preservar a ocorrência das palavras alvos no mesmo contexto prosódico e fonético, e evitar, dessa maneira, possíveis modificações nos parâmetros acústicos (como, por exemplo, a duração). Além disso, o uso de uma frase veículo teve como objetivo tentar controlar a curva entoacional da produção de cada palavra do corpus e, assim, evitar a curva ascendente característica da produção obtida por meio de repetição isolada em forma de lista. A frase-veículo é uma proteção que evita o efeito lista no contorno entoacional da palavra.

No caso das palavras nas quais as sequências C1+V ocupavam a posição pós-tônicas, foram utilizadas diferentes frases veículos, a saber: “digo ___ baixinho” e “digo ___ paciente”. Os resultados apresentados no próximo tópico permitiram-nos a configuração de nosso corpus final.

3.2.1 O estudo piloto: resumo dos resultados

Em línguas como o inglês, vogais não acentuadas são, frequentemente, mais desvozeadas do que vogais acentuadas, pois são mais breves e mais centralizadas, então mais inclinadas à redução. Este padrão costuma aparecer na maioria das

línguas naturais (DAVIDSON, 2006). Sendo assim, poderíamos prever que vogais não acentuadas seriam mais favoráveis ao desvozeamento, pois seus gestos são menores e mais facilmente sobrepostos por gestos de consoantes adjacentes. Essa hipótese mostrou-se, em parte, verdadeira para o português brasileiro, uma vez que todos os casos de vogais desvozeadas aconteceram na pós-tônica.

No teste piloto, todos os casos de desvozeamento ficaram restritos à posição pós-tônica. Segundo Gordon (1998), as posições não acentuadas e finais favorecem a sobreposição de gestos que provocam o desvozeamento. Diferentemente de outras línguas, como o Lezgi (CHITORAN, ISKAROUS, 2008), a posição pré-tônica não apresentou nenhum caso de desvozeamento nos contextos estudados.

Questão da altura vocálica

No que diz respeito à questão da altura vocálica, a principal condição fonética na qual a vogal torna-se desvozeada é similar: o desvozeamento ocorre principalmente com vogais altas. A figura abaixo mostra que há grande diferença entre as ocorrências de desvozeamento entre /i, u/ e /a/. Shiraishi (2003), em estudo restrito ao japonês, afirma que o desvozeamento de vogais médias e baixas é extremamente raro, restrito a poucas palavras, o que também foi atestado nos dados de nosso estudo piloto.

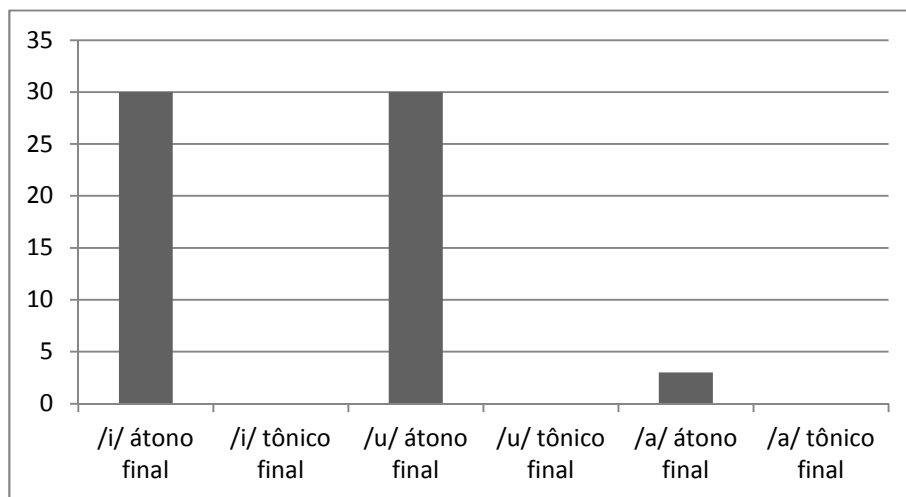


Figura 3: Frequência de ocorrência de desvozeamento por vogal no experimento piloto.

A co-ocorrência de desvozeamento total e parcial

Torreira e Ernestus (2010) já reportaram que a coexistência de vogais total e parcialmente desvozeadas é comum na análise do desvozeamento. Segundo os autores, a frequência de desvozeamentos parciais, isto é, aqueles que apresentavam uma porção de aperiodicidade do sinal vocálico (desvozeamento) e uma porção de periodicidade subsequente, foi maior quando a vogal era seguida por consoante sonora. Por outro lado, as vogais eram completamente desvozeadas, ou seja, totalmente aperiódicas e sem sinal de vozeamento, quando acompanhadas de consoantes surdas (cf. TORREIRA, ERNESTUS, 2010). Nos dados analisados em nosso estudo piloto, foi possível destacar a intensa co-ocorrência de ensurdecimento completo e parcial de vogais. Essa ocorrência também foi atestada em nosso corpus final e o fenômeno vai ser detalhado no Capítulo 4 da dissertação.

3.2.2 A montagem do corpus final

Uma vez que a análise do corpus piloto mostrou que somente as pós-tônicas favoreciam a ocorrência de vogais desvozeadas, um segundo corpus foi elaborado com o objetivo de explorar somente as pós-tônicas. Este segundo grupo de palavras (Anexo A), por sua vez, foi inserido em frases mais contextualizadas. Mesmo optando por um conjunto de frases que parecesse mais natural para nossos informantes, o rigor na composição das sentenças (controle do contexto consonantal e número de sílabas, por exemplo) permitiu um maior controle das variáveis. As palavras-alvo contêm as vogais /i, u, a/, tônicas e átonas (o ambiente de desvozeamento), em palavras onde C₂ é sempre a consoante [s]. Os seguintes ambientes foram comparados, como nos exemplos:

- [s] + [i], tônicas e átonas – ex. [sa'ʃi] / [ʃpase]

- [s] + [u], tônicas e átonas - [a'su] / [ʃasu]

- [s] + [a], tônicas e átonas: [i'sa] / [ʃkasa]

Multiplicando o número de sujeitos, pelo número de repetições e, finalmente, pelo o número de palavras, obtivemos um total de 180 ocorrências para serem analisadas.

5 (repetições) x 3 (vogais) x 2 (posições: átonas e tônicas) x 6 (sujeitos)= 180

3.3 Segmentação e rotulação dos dados

A segmentação dos dados foi realizada através da observação da forma da onda e do espectrograma. Primeiro, foi realizada uma inspeção visual dos dados, seguida pela divisão em arquivos separados de cada frase veículo produzida pelas falantes. Para cada arquivo foi criado um TextGrid (ferramenta de rotulação disponibilizada pelo software Praat que pode ser vista nas figuras abaixo). Depois da inspeção visual, foi realizada uma segmentação manual prévia e, em seguida, uma revisão minuciosa, que tentou minimizar erros. Feita a segmentação, foi realizada a rotulação dos dados, que possibilitou obtenção das medidas acústicas através de scripts. Os rótulos, conforme a Figura 4a abaixo, nos permitiam identificar (i) a palavra alvo, (ii) a sílaba tônica da palavra e (iii) a sílaba átona que caracteriza a fricativa + vogal desvozeada. No caso das palavras nas quais a vogal era vozeada, isto é, permitia a visualização da periodicidade, outro rótulo era adicionado com a identificação da vogal, conforme Figura 4b.

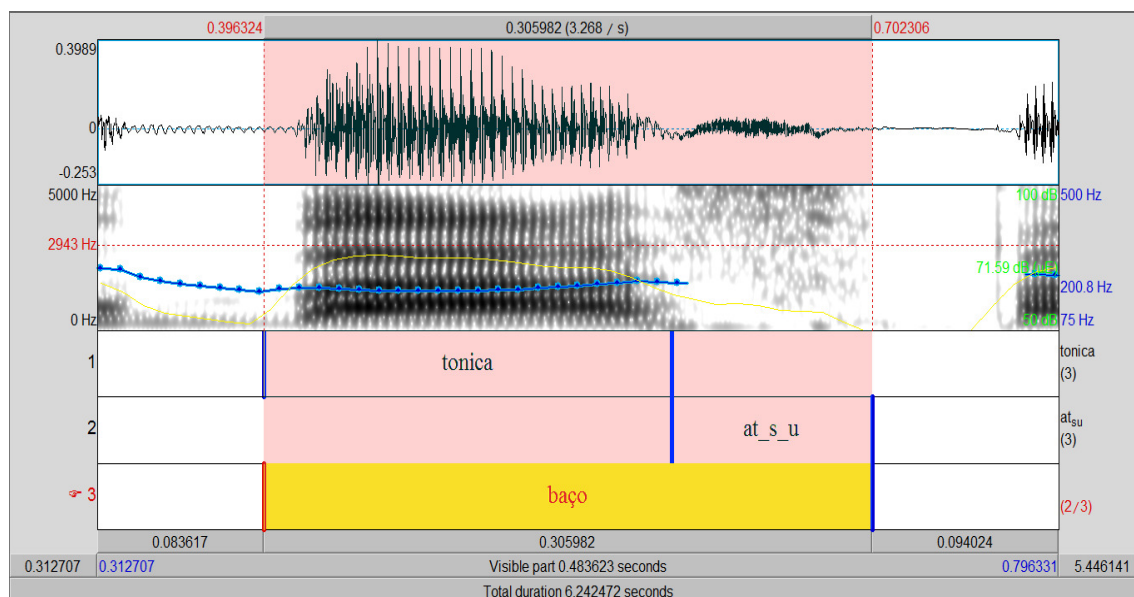


Figura 4a: Rótulos em *Text Grid* para dados referentes a sequencia fricativa + vogal desvozeada.

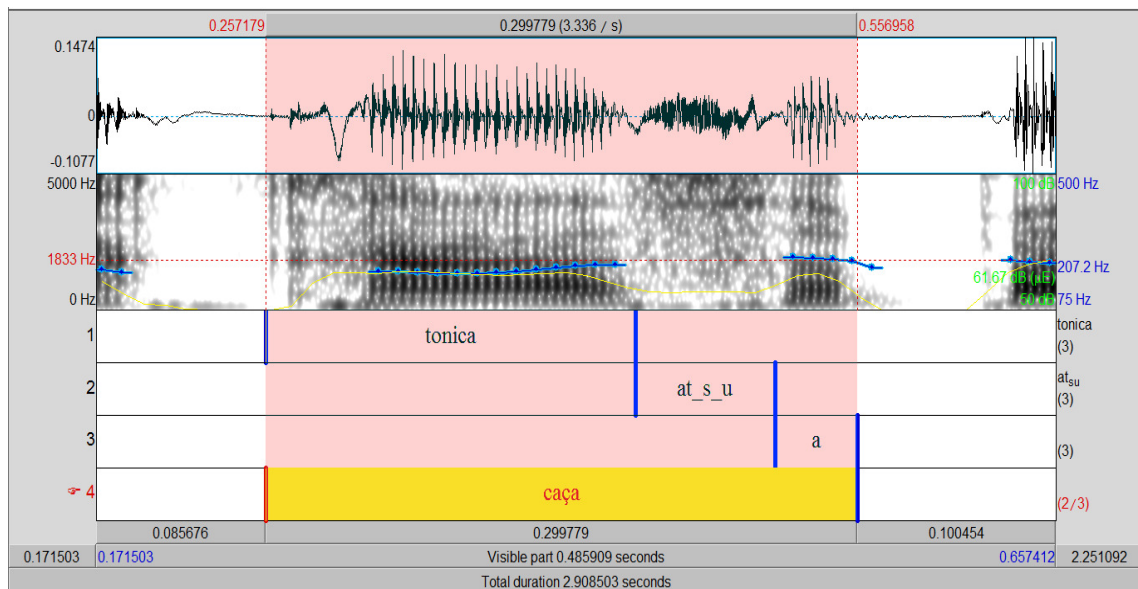


Figura 4b: Rótulos em *Text Grid* dos dados referente a sequencia fricativa + vogal vozeada.

3.4 Análise dos dados

Para a análise, realizamos os procedimentos básicos de edição e segmentação do programa *Praat for Windows* (v. 5.2.28). Depois da gravação, os sinais de fala foram armazenados diretamente em um HD a uma taxa de amostragem de 22KHz, a 16 bits. Cada repetição dos sujeitos constituiu um arquivo .wav para facilitar a organização dos arquivos. Por meio do *Praat*, separamos cada frase-alvo de nossa observação em um arquivo.

As medidas foram obtidas a partir da janela da forma de onda sincronizada ao respectivo espectrograma de banda larga, extraído por meio do software mencionado. Para a análise e identificação das vogais desvozeadas foi, em geral, observada a ausência de periodicidade no sinal acústico, como o exemplo da sílaba /su/ exibido abaixo:

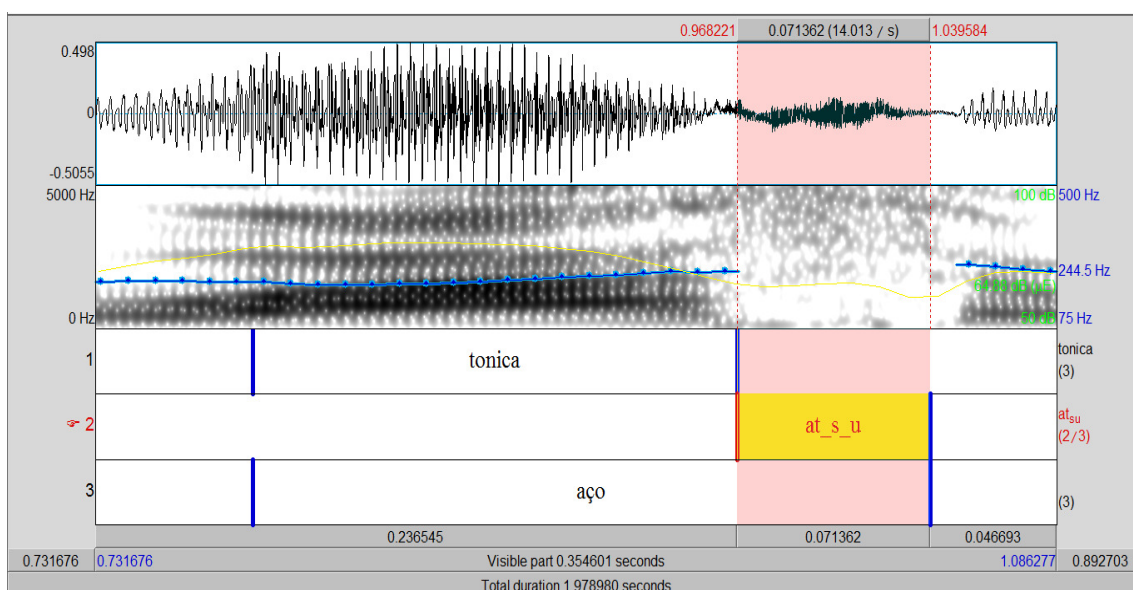


Figura 5: Ausência de periodicidade e de limites entre fricativa e vogal no sinal acústico.

A análise dos dados consistiu na medida da duração absoluta do ruído, quando as vogais eram desvozeadas, e da vogal quando estas foram vozeadas. As medidas e duração foram obtidas através do uso de um script. Além das medidas de duração, foi realizada a análise espectral de toda a porção do ruído que antecedia a vogal e da fricativa que precede a vogal não acentuada. As medidas foram realizadas automaticamente por um *script* que segue este procedimento: (i) gera uma janela FFT; (ii) calcula a duração do segmento rotulado e define seu centro; (iii) calcula as medidas dos quatro primeiros momentos espectrais da fricativa. Os dados são, então, transportados automaticamente para uma tabela.

3.4.1 Medidas de duração

As medidas de duração foram realizadas com duas diferentes unidades segmentais linguísticas: o ruído fricativo da consoante coronal surda /s/ e as vogais /i, u, a/ adjacentes às fricativas. É importante salientar, mais uma vez, que, nas palavras

nas quais as vogais eram desvozeadas, a duração do ruído representa a sequência fricativa + vogal desvozeada.

Vale ressaltar que todas as rotulações foram feitas com o auxílio visual da forma de onda juntamente com o espectrograma; e a duração foi obtida por meio do uso de script.

3.4.2 Medidas da Razão de Centralização Formântica e Área de Espaço Vocálico

A Razão de Centralização Formântica (FCR) é uma nova medida para a avaliação das distâncias relativas entre as vogais dos falantes (SAPIR *et al.*, 2010; KENT, KIM, 2003). Esse procedimento é empregado para normalizar os formantes vocálicos, e usado com sucesso para diferenciar a fala de disartricos da fala normal. A medida de FCR foi obtida pela fórmula³:

$$\text{FCR} = (\text{F2u} + \text{F2a} + \text{F1i} + \text{F1u}) / (\text{F2i} + \text{F1a}).$$

A observação da área de espaço vocálico (VSA), por sua vez, também pode dar pistas da atividade articulatória executada pelas informantes. Esses espaços foram obtidos pela fórmula:

$$\text{VSA} = \text{ABS} ((\text{F1i} * (\text{F2a} - \text{F2u}) + \text{F1a} * (\text{F2u} - \text{F2i}) + \text{F1u} * (\text{F2i} - \text{F2a})) / 2)$$

A avaliação do FCR aliada aos valores do VSA nos permite fazer inferências sobre a produção da fala, conforme SAPIR *et al.* (2010). Valores altos de FCR refletem centralização das vogais, enquanto que valores baixos indicam a sua expansão. Vogais centralizadas mostram possíveis efeitos coarticulatória, enquanto vogais periféricas indicam maior precisão de articulação. O inverso acontece com os valores de VSA, vogais mais centralizadas apresentam menores valores de VSA, obviamente.

³ Nas fórmulas de FCR e VSA, F1 e F2 são os valores referentes aos dois primeiros formantes de /a, i, u/. As medidas dos formantes foram feitas de forma manual através do Praat.

A observação do FCR e VSA permite fazer inferências sobre a atividade dos articuladores mandíbula, lábios e língua (LINDBLOM, SUNDBERG, 1971). Devido à sobreposição articulatória e possível centralização vocálica, o VSA pode ser reduzido e o FCR aumentado em vogais vozeadas que estão em posições de desvozeamento em relação às vogais de condições normais.

3.4.3 Análise da forma do espectro

Os momentos espectrais referem-se a uma métrica quantitativa baseada na análise estatística do espectro. Tais medidas conseguem evidenciar os gradientes de produção e carregam informações locais e da forma do espectro. Além disso, os momentos espectrais podem revelar informações quanto aos locais de constrição envolvidos na produção das fricativas. São calculadas, assim, a média, a variância, a assimetria e a curtose da distribuição dos componentes do espectro:

- ***Centróide/Primeiro Momento***: o centróide (ou centro de gravidade) é a média das frequências ponderadas pela intensidade apontando para o centro de gravidade dado pelo espectro do ruído da fricativa (FFT).

Para a análise espectral, seguimos a metodologia proposta por Forrest *et al* (1988). As medidas dos momentos espectrais foram obtidas a partir de espectros FFT (*Fast Fourier Transform*). As medidas são obtidas através da janela de objetos do Praat, selecionando a opção *Spectrum slice* e extraindo automaticamente os valores dos momentos a partir da *Query* na janela de opções do Praat, como pode ser visto na figura abaixo⁴:

⁴ Para uma explicação mais detalhada da obtenção dos momentos espectrais por meio do *Praat*, ver Berti (2006).

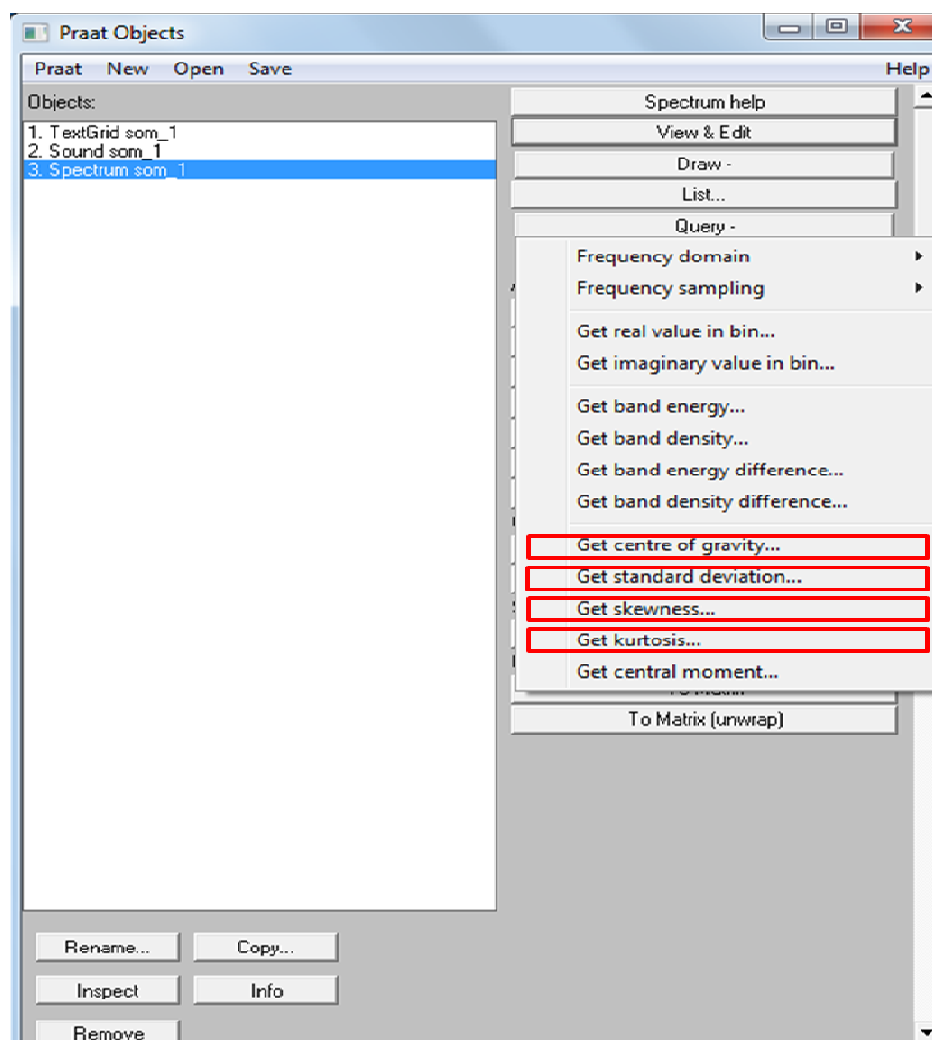


Figura 6: Janela de objetos do Praat para a obtenção do valor referente ao centróide, desvio padrão, assimetria e curtose.

3.5 *Análise Estatística*

Para a análise estatística, utilizamos um Teste T pareado através do qual foram comparados os valores de centróide obtidos do sinal das fricativas e os valores de FCR e VSA. Ainda para o FCR e VSA, um teste de correlação de Pearson foi usado para verificar a relação entre as duas medidas. Para as medidas de duração foi utilizado o teste de Análise de Variância, depois, para verificar as diferenças entre

pares de médias, aplicou-se o teste de Tukey. As análises foram feitas pelo software *Statistics*, v. 7.

3.6 O uso de Scripts

A análise acústica é um processo que deve ser realizado com extremo cuidado para evitar erros de medida. Para tanto, durante as medições, optamos por realizar semi-automatizações dessas análises. Em anexo seguem os scripts utilizados na íntegra.

Neste trabalho utilizamos dois scripts⁵: o primeiro realiza as medidas dos quatro momentos espectrais de todo ruído; o segundo script foi utilizado para obtenção das medidas de duração. Ambos estão reproduzidos nos anexos da dissertação e foram obtidos por meio da dissertação de mestrado de Rinaldi (2010)

⁵ Os scripts foram alterados de acordo com os objetivos de nosso estudo. Por exemplo, no script que mensurava o centroide do ruído foram acrescentadas duas funções de filtragem: *stop band* 500hz e *pass band* 4Khz a 9Khz. O objetivo da filtragem será explicado no capítulo 5.

Capítulo 4

Resultados e discussão I: Descrição dos dados

4.1 Considerações Preliminares

Esta seção, que traz os resultados de nosso trabalho, está organizada em duas partes: na primeira, descrevemos os dados, mostrando a ocorrência de semelhanças e as diferenças entre os informantes; na segunda, discutimos os dados de nossa análise acústica. Dentro de cada subseção vamos discutir vários aspectos em torno do processo de desvozeamento.

Recapitulando, foram gravados seis informantes que leram as frases-veículos com cinco repetições. As gravações foram segmentadas de acordo com o procedimento descrito na Metodologia (Capítulo 3 da dissertação). Os dados foram extraídos das segmentações de acordo com as variáveis descritas anteriormente.

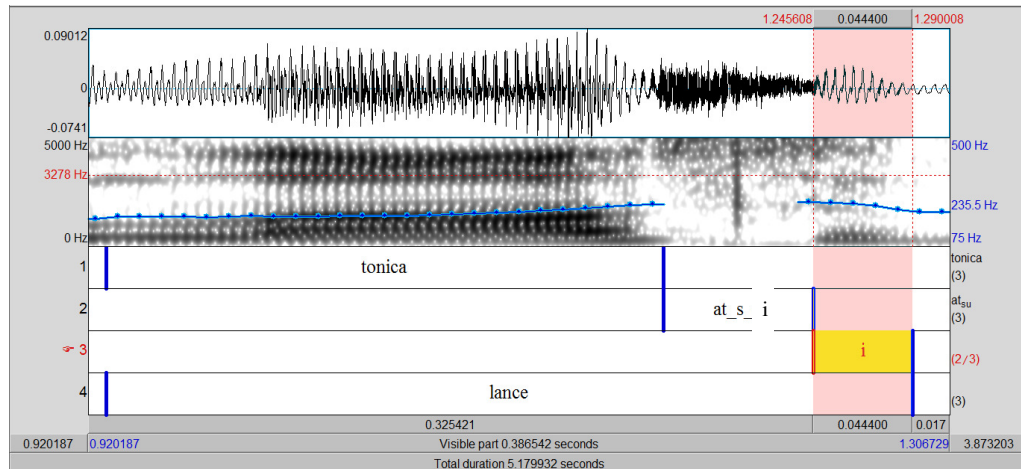
Um dos objetivos deste estudo foi determinar as semelhanças e possíveis diferenças do desvozeamento vocálico no PB com a ocorrência do fenômeno em

algumas outras línguas. Como esperado, todas as informantes produziram vogais desvozeadas. No entanto, houve mais variação do que o esperado no padrão de ocorrência de desvozeamento entre os sujeitos. Este achado, por sua vez, se distancia dos resultados que a literatura reporta, como veremos no tópico a seguir.

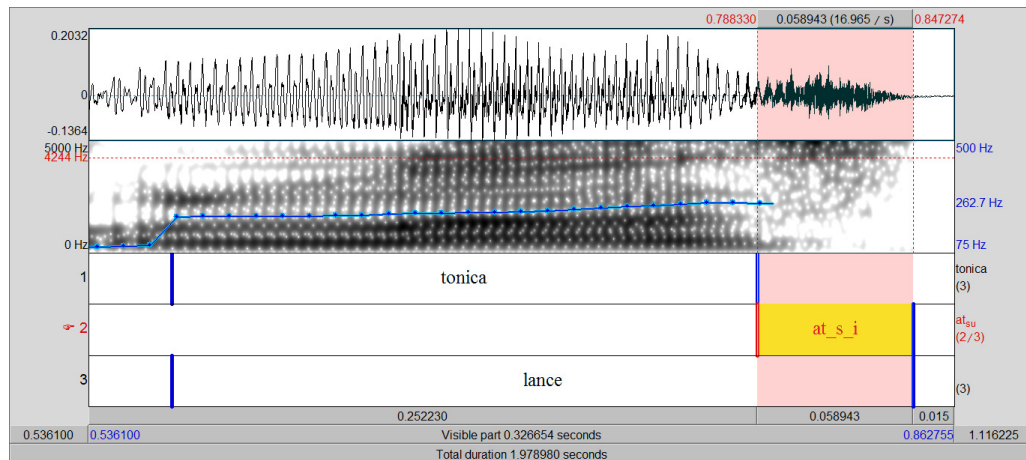
4.2 Características das vogais desvozeadas do PB

Os dados analisados neste tópico são os de dissílabos com estrutura [´C₁VC₂V]. A sílaba átona é formada pelas vogais /i, u, a/ e a C₂ é sempre a consoante [s]. As palavras foram ditas em uma frase-veículo contextualizada, como por exemplo: “O passe parecia ir em direção ao gol”.

Em geral, no desvozeamento vocálico, as vogais altas /i/ e /u/ perdem sua sonoridade, percebida articulatoriamente pela falta de vibração das pregas vocais e acusticamente pela ausência do sinal de frequência fundamental no sinal acústico (cf. KONDO 2005, MAEKAWA, KIKUCHI, 2005). Mais especificamente, o sinal acústico da vogal normalmente carece dos componentes que são esperados para a vogal vozeada, a saber, uma forma de onda periódica regular, a barra de vozeamento e uma configuração formântica bem definida (CF. HAN 1962, BECKMAN, SHOJI 1984, KONDO 2005, TSUCHIDA 1997). A ausência desses componentes é ilustrada na Figura 7, onde comparamos um /i/ não afetado pelo desvozeamento em (a) com /i/ desvozeado em (b).



(a)



(b)

Figura 7: Exemplo de /i/ vozeado em (a) e /i/ desvozeado em (b) para a palavra “lance” /laNsi/. As porções destacadas indicam em (a) periodicidade da vogal vozeada e em (b) a extensão do desvozeamento total.

A presença de estrutura formântica no caso de desvozeamento não está clara em (b), como também já foi evidenciado na literatura (por exemplo, HAN 1962). De acordo com Kondo (2005), a realização acústica de vogais desvozeadas é a continuação da consoante anterior se a consoante precedente é uma fricativa. Esta característica pode ser observada no espectrograma de 7b, comparada aos limites bem definidos entre fricativa e vogal, em 7a.

Vale ressaltar que nem todas as produções de nossos sujeitos se encontravam nos padrões das figuras acima. Com relação ao desvozeamento, é bem comum relatos de produções gradientes em termos de grau. Por exemplo, Beckman (1996) e Kondo (1997) encontraram ensurdecimento em vários graus. Torreira e Ernestus (2010), em uma investigação sobre o desvozeamento no Francês, reportam a coexistência entre vogais totalmente desvozeadas, como visto em 7b, e desvozeamentos parciais, como encontrado na Figura 8. Em comparação ao desvozeamento total, as ocorrências de vozeamento parcial contam com extensa fricção e barra de vozeamento curta. Já em comparação às produções vozeadas, nos casos que aqui designamos de desvozeamento parcial há ruído sobreposto à periodicidade da vogal.

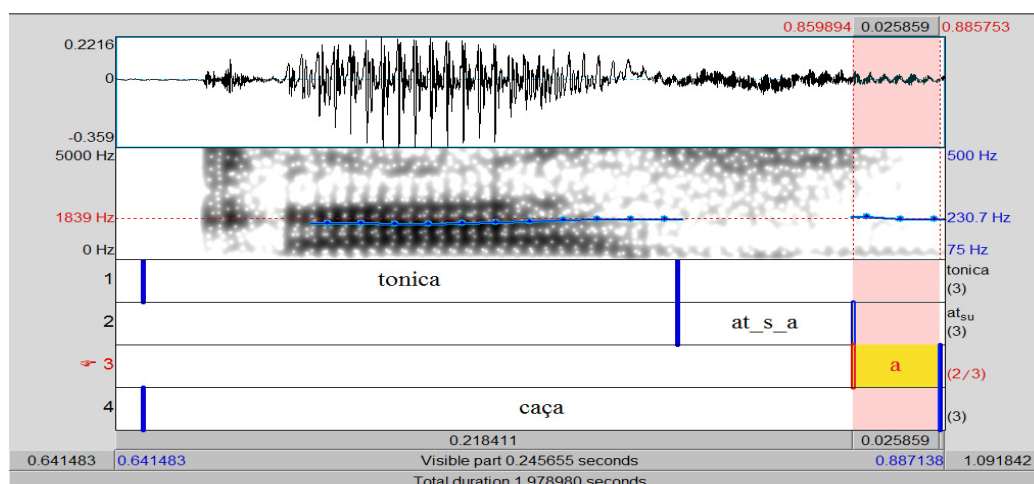


Figura 8: Exemplo de /a/ parcialmente desvozeado para a palavra “caça” /ˈkasa/.

Em geral, os dados observados neste experimento foram semelhantes aos relatados na literatura, isto é, grande variabilidade entre as três produções possíveis para as vogais. Para Tsuchida (1997) o fenômeno do desvozeamento é gradiente porque algumas vogais são parcialmente ou totalmente vozeadas. Para o francês, por exemplo, somente um terço das vogais-alvos é completamente desvozeado (TORREIRA, ERNERTUS, 2010). Nos dados analisados aqui, também é possível

destacar a ocorrência de ensurdecimento completo e parcial de vogais e de vogais vozeadas.

Para efeito de uma comparação mais geral, a figura abaixo mostra que, em número total de ocorrências, a produção de vogais vozeadas e desvozeadas são equivalentes.

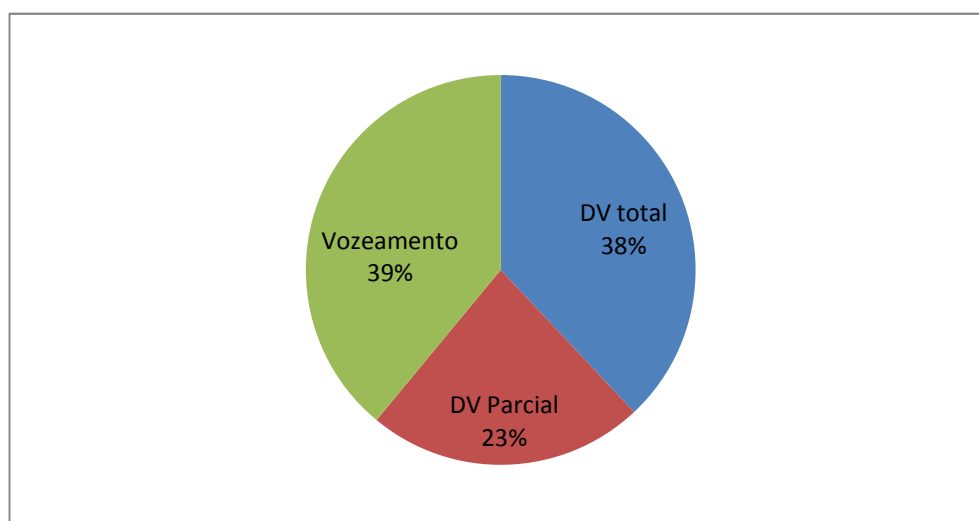


Figura 9: Ocorrências de desvozeamento total (DV total), desvozeamento parcial (DV Parcial) e de vogais vozeadas em todo conjunto de dados.

É importante ressaltar que grande parte dos 39% de produções de vogais vozeadas é da vogal baixa. Com exceção dos dados de Smith (2002), por exemplo, nenhuma ocorrência de ensurdecimento com vogais baixas foi relatada na literatura, mas, nos dados deste trabalho, isso ocorreu, embora muito menos frequentemente do que com vogais altas. A ocorrência de desvozeamento nas vogais altas também variou bastante entre os sujeitos. A Figura 10 abaixo mostra a frequência com que ensurdecimento total e parcial e a produção de vogais vozeadas ocorreram nas três vogais testadas.

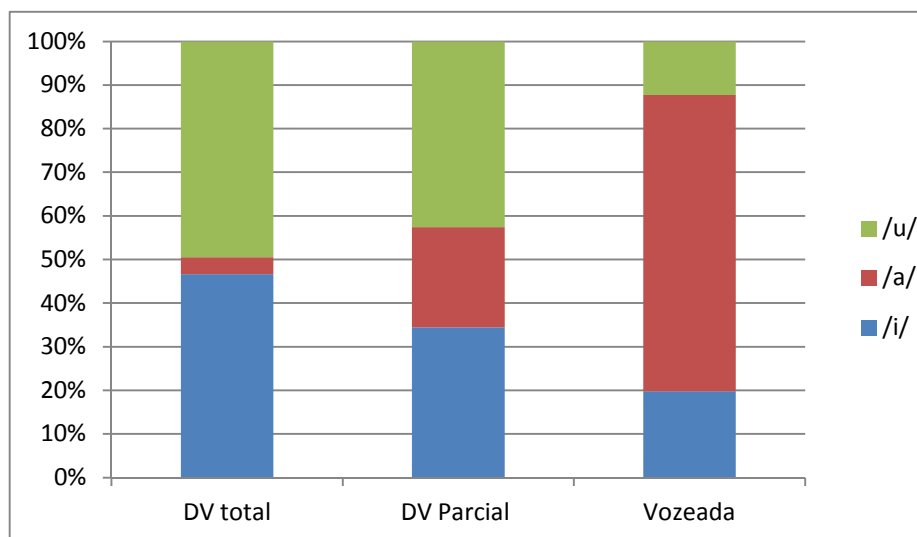


Figura 10: Taxa de desvozeamento total (DV total), desvozeamento parcial (DV Parcial) e vogal vozeada para /u/, /a/ e /i/.

Fica claro na leitura do gráfico acima que a vogal /a/ foi responsável pela maior parte das vogais vozeadas, enquanto /i/ e /u/ dividiram a grande maioria das produções de vogais desvozeadas. Recorrente neste trabalho e consenso na literatura é o fato de que o desvozeamento ocorre principalmente com vogais altas em ambiente de segmentos surdos (JUN, BECKMAN, 1993).

Shiraishi (2003), em estudo restrito ao japonês, afirma que o desvozeamento de vogais médias e baixas, assim como o desvozeamento de vogais acompanhadas de consoantes vozeadas (YOSHIDA, SAGISAKA, 1990), é extremamente raro e restrito a poucas palavras.

É importante ressaltar que apenas três das informantes produziram a vogal baixa desvozeada. Em contraste, todos os seis sujeitos produziram tanto /i/ quanto /u/ desvozeado, embora o número de produções tenha variado entre as falantes. A variabilidade entre os informantes pode ser atestada nos gráficos abaixo:

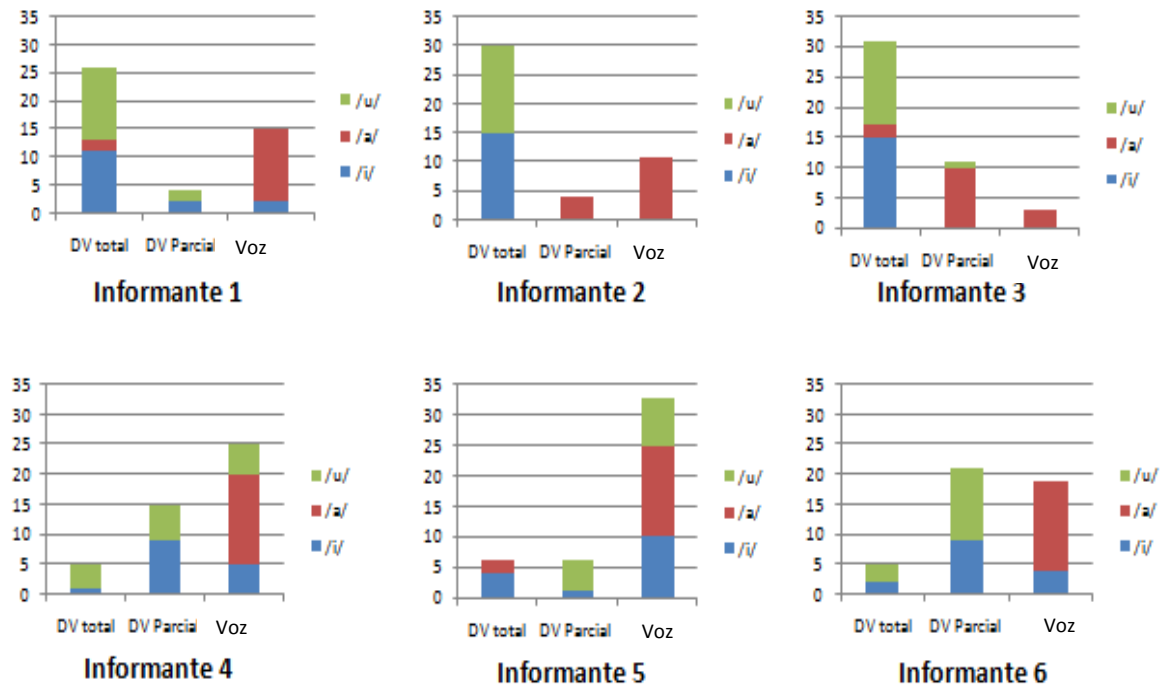


Figura 11: Taxa de realização de desvozeamento total (DV total), desvozeamento parcial (DVParcial) e vogais vozeadas (voz) para cada sujeito analisado.

Os sujeitos 1, 2 e 3 são aqueles que mais produzem vogais desvozeadas. O sujeito 6 também tem as vogais altas afetadas por efeitos de desvozeamento, no entanto suas produções são, em maior número, de vogais desvozeadas parciais. As informantes 4 e 5 também produzem vogais desvozeadas, mas em número bem reduzido em relação aos primeiros.

A variabilidade de produção de vogais totalmente desvozeadas entre os sujeitos pode ser melhor visualizada no gráfico abaixo. É importante lembrar que os sujeitos desta pesquisa eram da mesma cidade, sexo, faixa etária e aproximadamente o mesmo nível sócio econômico, o que neutraliza, ao menos em parte, a influência de fatores sociolinguísticos.

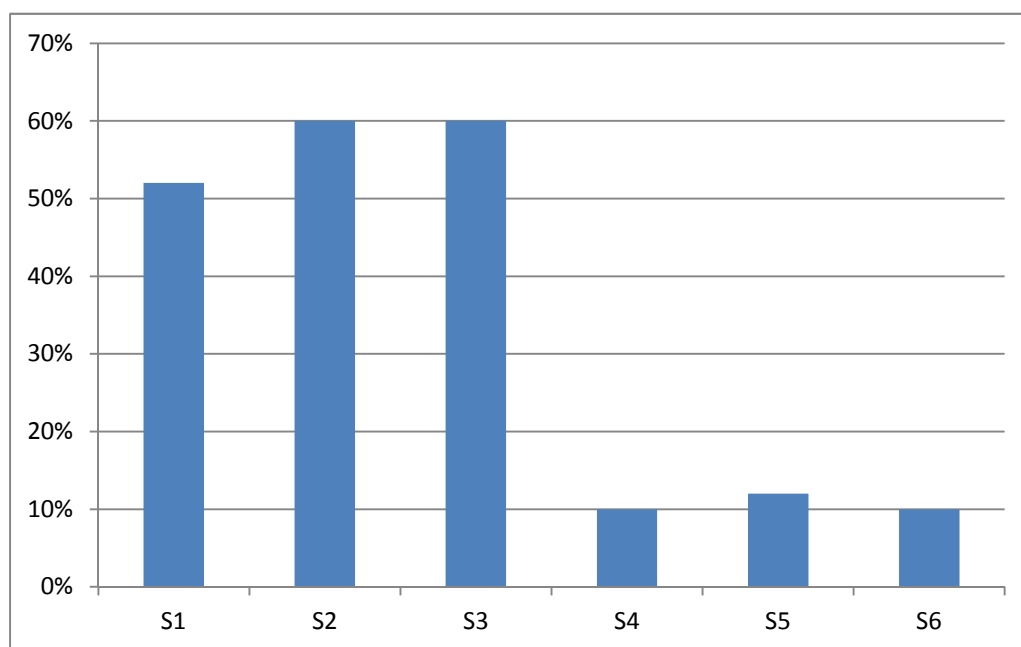


Figura 12: Percentagem de vogais desvozeadas por sujeito.

Outro fator que merece destaque é a influência ou não da frequência de ocorrência das palavras sobre a ocorrência de vogais desvozeadas. Mesmo não sendo um objetivo deste trabalho, verificar se a frequência de ocorrência das palavras no léxico influi na produção de vogais desvozeadas pode ser uma questão para futuras investigações. A frequência de ocorrência das palavras no léxico e suas consequências vêm sendo amplamente estudadas. Aqui, parece não haver efeitos de frequência sobre a variação de produção de vogais surdas:

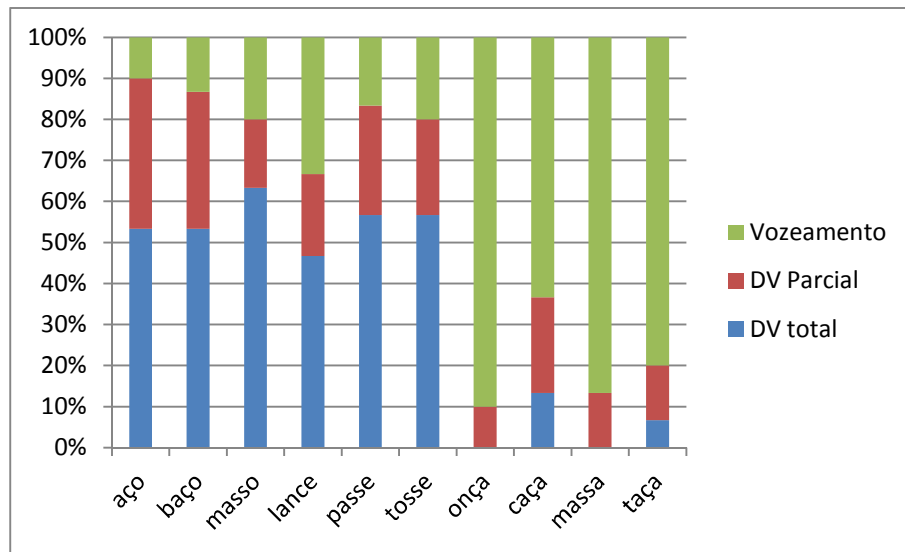


Figura 13: Taxa de desvozeamento total (DV total), desvozeamento parcial (DV Parcial) e vogal vozeada para cada palavra.

Tabela 1: Frequência de ocorrência das palavras alvo no PB.⁶

<i>Palavra-alvo</i>	<i>Transcrição</i>	<i>freq.</i>
passe	'pasl	1631
lance	'laNsI	1019
taça	'tasa	786
caça	'kasa	743
aço	'asU	703
massa	'masa	290
onça	'oNsA	217
tosse	'to'sI	132
masso	'masU	94
baço	'basU	52

⁶ As estimativas de frequência foram retiradas do banco de dados do CETEN-Folha. A base de dados do português brasileiro do CETEN tem corpus retirado do jornal Folha de São Paulo. Base elaborada pelo Núcleo Interinstitucional de Linguística Computacional (NILC), disponível em: <http://www.linguateca.pt/cetenfolha>.

Nota-se, analisando Tabela 1 e a Figura 13 acima, que, tanto palavras com maior (ex.: passe e lance) e menor (ex.: masso e baço) frequência na língua sofrem desvozeamento vocálico. O teste de correlação de Spearman, apesar de apresentar um valor relativamente alto (0.80), não foi significativo. Assim, não se pode afirmar que há relação direta entre a frequência e a percentagem de desvozeamento, mas não se pode descartar a possibilidade. A questão deve, portanto, ser investigada em pesquisas futuras com uma amostra maior e o controle apropriado.

4.3 Discussão

Inúmeras análises de línguas naturais reportam a ocorrência de vogais desvozeadas em diferentes condições (ANDREEVA, KOREMAN, 2008; KONDO, 1997; TSUCHIDA 1997, JUN, BECKMAN, 1993).

Em geral, os resultados relatados na literatura se repetiram nos dados descritos em nossa investigação. As vogais altas, que são intrinsecamente breves, muitas vezes se encontram desvozeadas em ambientes consonantais surdos. Além disso, merece destaque:

- (i) a ocorrência, mesmo que em menor escala, de vogais baixas desvozeadas;
- (ii) a intensa variação entre falantes;
- (iii) a ocorrência de vogais totalmente desvozeadas e parcialmente desvozeadas.

Como já foi insistentemente relatado em nosso trabalho, o desvozeamento tem sido entendido como um fenômeno relacionado a uma sobreposição de gestos. Dentre os resultados descritos até aqui, aquele que mais advoga a favor de um efeito de sobreposição de gestos, em detrimento de síncope vocálica, é a ocorrência de

desvozeamento total e parcial nos dados analisados. Se os desvozeamentos, completo e parcial, ocorrem em contextos semelhantes e, em alguns casos, para um mesmo sujeito, devem ser condicionados pelos mesmos fatores e o limite entre a ocorrência de um ou do outro parece mínimo. Aparentemente, um maior ou menor grau de sobreposição pode implicar na vibração mínima ou na não vibração das pregas vocais. Esta observação sugere que uma pista da vogal está presente na articulação, embora as pregas vocais possam ou não vibrar.

Além disso, o fato de as ocorrências variarem entre três possibilidades, a saber, vozeamento completo, desvozeamento completo ou desvozeamento parcial, mostra que a mudança não é categórica [+ voz] para [- voz], mas um processo gradual.

Mais pistas acústicas e articulatórias podem evidenciar ou não a hipótese de desvozeamento em detrimento da queda vocálica. Pesquisas anteriores sobre vários idiomas mostraram que ensurdecimento da vogal é geralmente acompanhado por uma compressão temporal da vogal e suas consoantes adjacentes. Nós verificamos que isso também é verdade para o PB procurando relações estatísticas entre a duração da sílaba e a ocorrência de vogais vozeadas e de desvozeamento total e parcial. No próximo capítulo, discutiremos a análise de dados de duração, as medidas do centroide e considerações sobre o FCR e VSA das vogais vozeadas e parcialmente desvozeadas.

Capítulo 5

Resultados e discussões II: Análise inferencial

5.1 Queda ou sobreposição de gestos: a análise da duração

Já há algum tempo pesquisas têm mostrado que o ensurdecimento da vogal pode ocorrer como resultado do aumento de sobreposição entre os gestos vocálico e consonantal. Sob essa hipótese, pode-se esperar que as vogais não vozeadas exibam sinais de aumento de coarticulação na sequência C+V.

Nesta seção vamos considerar a duração de segmentos e as possíveis pistas que esse parâmetro pode trazer para a discussão em torno do desvozeamento de vogais. Pesquisas anteriores sobre vários idiomas mostraram que ensurdecimento da vogal é geralmente acompanhado por uma compressão temporal da vogal e expansão de consoantes circunvizinhas (SMITH, 2002; FAGYAL AND MOISSET, 1999), resultante de um efeito coarticulatório. Torreira e Ernestus (2010) afirmam que as vogais surdas do Francês, por exemplo, são comprimidas e desvozeadas. Estudos

baseados no japonês demonstraram que a duração de sílabas desvozeadas é geralmente menor do que de sílabas vozeadas (TSUCHIDA, 1997).

As sílabas comparadas aqui foram as átonas finais das palavras do corpus, como, por exemplo, *aço para /u/* e *lance para /i/*. A duração dessas sílabas foi comparada quando as vogais eram ou não desvozeadas. As figuras 14 e 15, referentes à duração das sílabas átonas, permitem observar a diferença de duração existente entre as sílabas desvozeadas, parcialmente desvozeadas e sílabas com vogais vozeadas para /u/ e /i/ respectivamente.

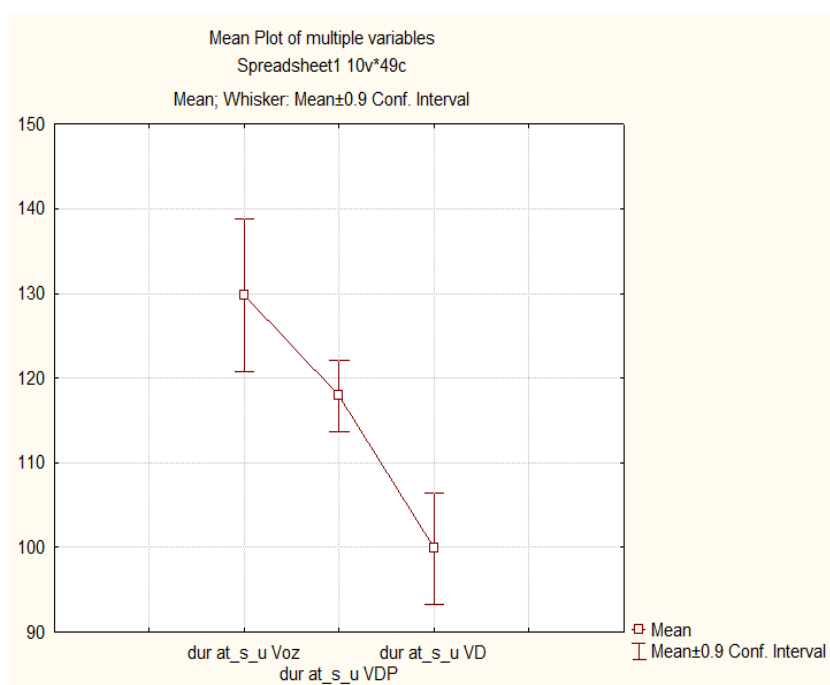


Figura 14: Duração (ms) das sílabas /s/ + vogal vozeada (at_s_u Voz), /s/ + vogais desvozeadas parcial (at_s_u VDP) e /s/ + vogal desvozeada total (at_s_u VD) para a vogal /u/.

Tabela 2: Análise de variância para duração das sílabas /s/ + vogal /u/ vozeada, /s/ + vogal /u/ desvozeada parcial e /s/ + vogal /u/ desvozeada total.

	Média (ms)	Grau de Liberdade	(H)	P
dur at_s_u Voz	129			
dur at_s_u VDP	117	2	18	0,0001
dur at_s_u VD	99			

Observando a Figura 14, verifica-se claramente que há diferenças entre as médias de duração para sílabas vozeadas em relação ao segundo grupo de médias, das sílabas parcialmente desvozeadas, e em relação à duração das sílabas com vogais desvozeadas. A Tabela 2 mostra o valor de p significativo para a diferença entre as médias. O mesmo padrão é observado na Figura 15, referente à média de sílabas compostas por /i/.

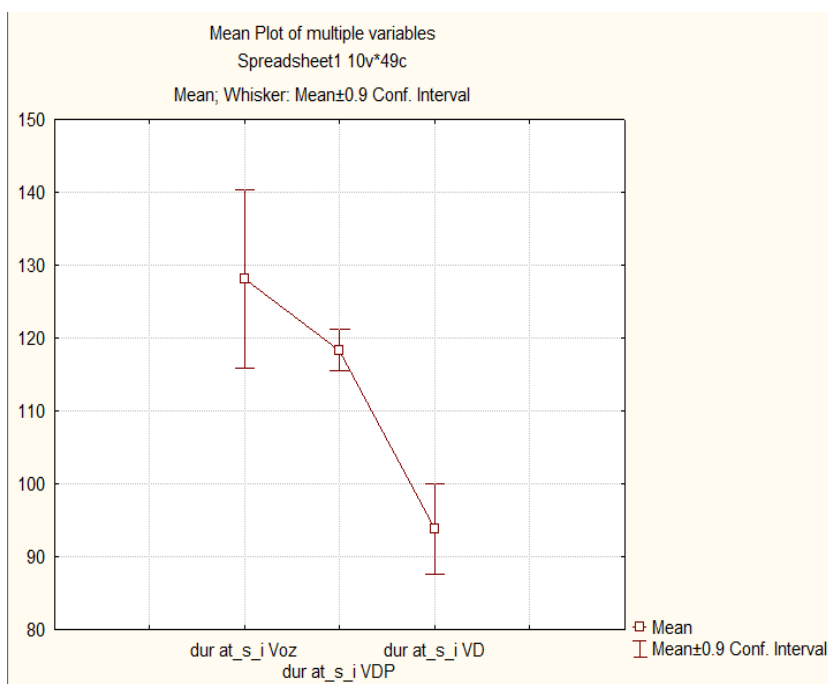


Figura 15: Duração (ms) das sílabas /s/ + vogal vozeada (at_s_i Voz), /s/ + vogais desvozeada parcial (at_s_i VDP) e /s/ + vogal desvozeada total (at_s_i VD) para a vogal /u/.

Tabela 3: Análise de variância para duração das sílabas /s/ + vogal /i/ vozeada, /s/ + vogal /i/ desvozeada parcial e /s/ + vogal /i/ desvozeada total.

	Média (ms)	Grau de Liberdade	(H)	P
dur at_s_i Voz	128			
dur at_s_i VDP	118	2	30,02	0,0001
dur at_s_i VD	93			

Como pode ser observado nas figuras e tabelas acima, a duração das sílabas desvozeadas foi menor se comparada com sílabas vozeadas ou parcialmente desvozeadas, tanto para /si/ quanto para /su/. Essa “compressão” da sílaba desvozeada, enquanto efeito coarticulatório, corrobora resultados de trabalhos como os citados acima. No entanto, a redução também pode ser interpretada com um falso efeito da queda da vogal em detrimento da hipótese de sobreposição de gestos, uma vez que o ruído aperiódico característico do desvozeamento tem duração reduzida em comparação a uma sílaba naturalmente vozeada.

Beckman (1982) compara a duração de sílabas CV, cuja vogal sofre desvozeamento, e sílabas CV em que a vogal é vozeada. Segundo a autora, o termo ensurdecimento é mais correto do que exclusão psicologicamente, mas fisicamente o termo "exclusão" caracteriza melhor o ensurdecimento, pois geralmente não há evidência acústica de existência da vogal. De acordo com a autora, as medidas de duração normalmente não mostram sinais de existência de uma vogal. No entanto, a ausência de evidência acústica óbvia não implica que o gesto articulatório tenha desaparecido inteiramente.

Para se excluir a hipótese de apagamento, uma segunda medida de duração foi realizada. No lugar da sílaba, somente a duração das fricativas foi considerada. Assim, se a duração das fricativas fosse igual, a hipótese de apagamento seria reforçada. Por outro lado, se a vogal ainda estiver presente na produção, enquanto

vogal desvozeada, a fricativa⁷ deve ser mais longa em sílabas desvozeadas comparada a uma fricativa em sílaba com vogal presente. Essa hipótese foi testada também por meio de Análise de Variância:

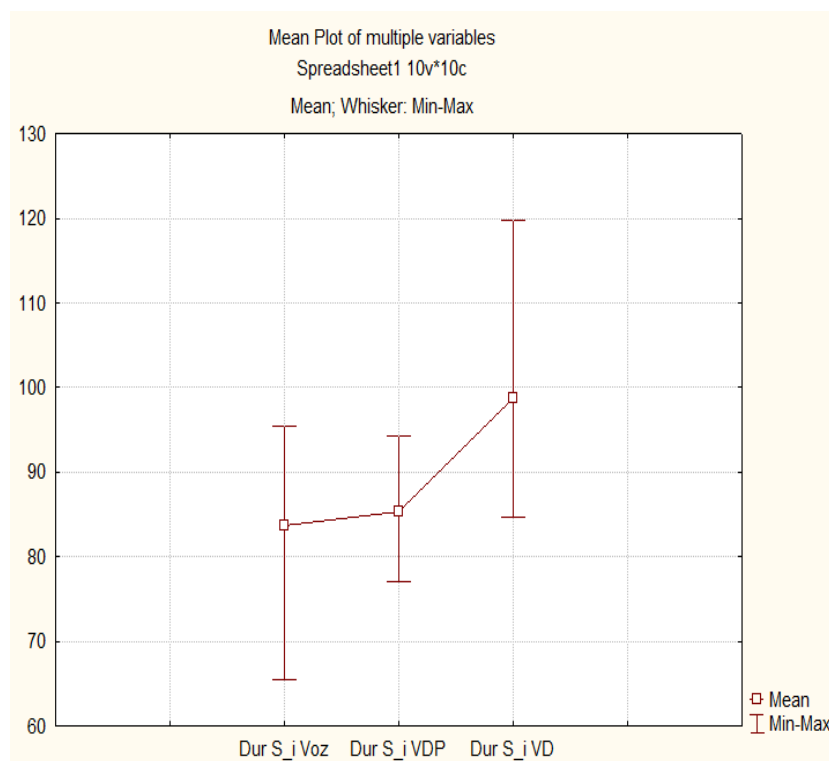


Figura 16: Duração (ms) do /s/ que acompanha vogal /i/ vozeada (S_i_Voz), parcialmente desvozeada (S_i_VDP) e totalmente desvozeada (S_i_VD).

Tabela 4: Análise de variância para a duração do /s/ que acompanha vogal /i/ vozeada, parcialmente desvozeada e totalmente desvozeada.

	Média (ms)	Grau de Liberdade	(H)	P
Dur S_i Voz	84	2	14,18	0,0008
Dur S_i VDP	85			
Dur S_i VP	98			

⁷ A duração da fricativa se faz ainda mais importante, uma vez que a vogal desvozeada perde periodicidade no sinal e é impossível distinguir a fricativa da vogal.

Observando a Figura 16, ao contrário do padrão observado para a comparação entre sílabas, verifica-se a duração da fricativa é menor em sílabas com vogal /i/ vozeada. A duração da fricativa em ambiente desvozeado é maior. As diferenças entre as médias também foram altamente significativas, como mostra a Tabela 4.

Para /u/, o padrão mantém-se o mesmo: a duração da fricativa em ambiente de /u/ desvozeado é maior do que a duração das fricativas de /u/ parcialmente desvozeado e vozeado, como se observa na figura 17 e na Tabela 5.

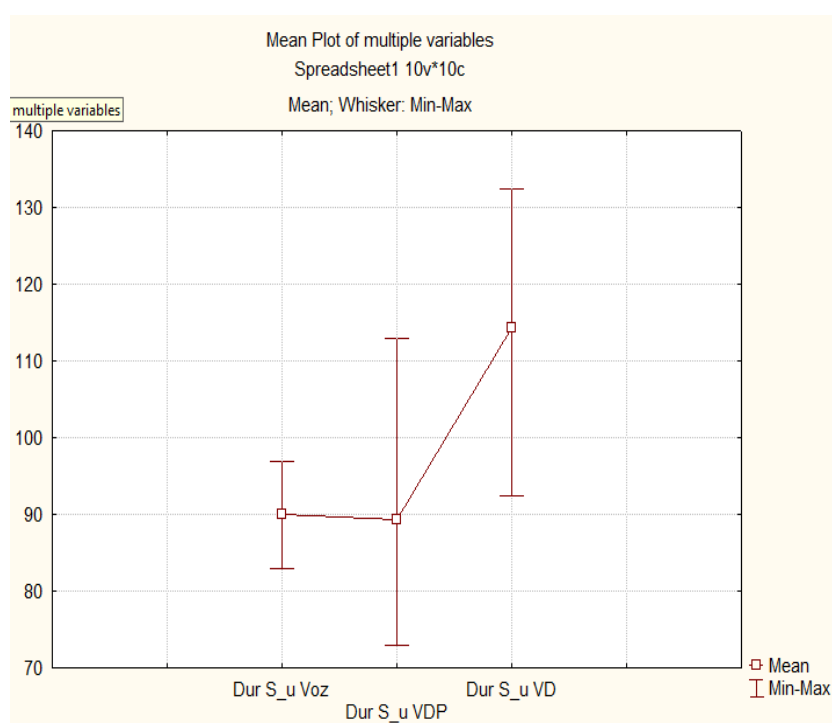


Figura 17: Duração (ms) do /s/ que acompanha vogal /u/ vozeada (S_u_Voz), parcialmente desvozeada (S_u_VDP) e totalmente desvozeada (S_u_VD).

Tabela 5: Análise de variância para duração do /s/ que acompanha vogal /u/ vozeada, parcialmente desvozeada e totalmente desvozeada.

	Média (ms)	Grau de Liberdade	(H)	P
<i>Dur S_u Voz</i>	90			
<i>Dur S_u VDP</i>	89	2	9,621	0,0081
<i>Dur S_u VP</i>	114			

Os dados acústicos exibidos nas Figuras 16 e 17 são consistentes com uma hipótese de sobreposição gestual e não são facilmente explicados por uma análise de exclusão, como sugere Beckman (1982). A duração do /s/ é significativamente maior quando a vogal - /i/ para a Figura 14 e /u/ para a Figura 17 - é desvozeada. Este achado é consistente com uma situação em que o ruído da fricativa e o gesto de abertura glótica se combinam encobrendo totalmente a vogal, fazendo parecer que a vogal tivesse sido excluída. O dado é semelhante aos resultados para o /s/ do francês (SMITH, 2002) e consistente com a possibilidade de que a vogal seja de fato produzida, mas sobreposta pela consoante precedente.

Se a vogal realmente fosse excluída, seria esperado que o /s/ não apresentasse diferenças de duração nas três condições analisadas, a saber: fricativa acompanhada de vogal vozeada, parcialmente desvozeada e totalmente desvozeada. Parece, até este ponto, que desvozeamento é acompanhado ou acarretado pela redução temporal de apenas a parte oral da sílaba, em vez de a sílaba como um todo. O aumento observado na coarticulação CV em casos de desvozeamento, como os dados de duração mostram, indica que a sobreposição articulatória de C sobre V pode ser a principal causa deste encurtamento da vogal, resultando no desvozeamento vocálico.

Em termos de duração da vogal, sílabas CV são significativamente menores quando a vogal sofre desvozeamento do que quando elas não sofrem

desvozeamento. A vogal parece pouco presente na dimensão temporal. No entanto, apesar desse quase desaparecimento da vogal, a articulação da vogal parece permanecer como “rastro”. Com respeito à acústica e a pista articulatória da vogal, os dados de duração evidenciam sem complicação: a vogal é surda; e, assim mesmo, se as características de vozeamento desaparecerem, os correlatos acústicos da presença da vogal permanecem.

5.2 Pistas acústicas: a análise do centróide

As palavras deste estudo foram selecionadas com o objetivo de favorecer a produção de vogais desvozeadas. Como já foi dito, as vogais-alvos estavam sempre na posição átona e acompanhadas da fricativa alveolar surda. Esse controle, no entanto, não impediu a grande variabilidade de produção das vogais. Assim, foi possível perceber alguma forma de instabilidade ou falta de constância na produção dos falantes que escolhemos como sujeitos desta pesquisa. Essa variabilidade já foi descrita e pôde ser percebida na Figura 9 do tópico 3.2.1 de nossos resultados.

Enquanto S1, S2 e S3 produzem maior número de vogais desvozeadas, S4, S5 e S6 oscilam mais, embora cheguem a produzir um bom número de vogais desvozeadas. Apesar da semelhança, não foi possível encontrar nenhum padrão comparativo entre os grupos. Essa falta de padrão dificultou nosso objetivo de determinar as diferenças entre os padrões do primeiro momento espectral das vogais, o centróide.

Como se sabe, alguns estudos mostraram que o centróide é um parâmetro importante no reconhecimento da qualidade vocálica (CHISTOVICH, LUBLINSKAYA, 1979, CHISTOVICH, 1979). O centróide ou primeiro momento da distribuição espectral é uma medida estatística que indica a concentração de energia (FORREST *et al*, 1988). Chistovich (1979), no seu estudo, reporta que vogais sintetizadas com

baixo centro de gravidade foram, quase sempre, identificadas como /u/, mostrando que o centróide é um parâmetro importante no reconhecimento dos falantes.

Para Tsuchida (1994) o centróide da região fricativa é uma pista de extremo valor na identificação das vogais. Para ela, as frequências médias de [ʃ] foram sempre superiores às de [u] para ambas as sílabas de vogais surdas e sonoras. Conforme a autora, este dado é um efeito da coarticulação vogal seguinte. Além disso, Tsuchida (1994) mostra que, considerando toda a extensão da vogal desvozeada, o centróide de [u] desvozeado foi menor do que [u] vozeado. Entretanto, [ʃ] desvozeado e vozeado não apresentaram diferenças significativas. Em conclusão, a autora afirma que centros de gravidade mais baixos fornecem pistas de presença de vogal.

Uma análise como a de Tsuchida (1994) só foi possível para três sujeitos deste estudo, pois tanto S1 quanto S5 e S6 produziram /si/ com vogal desvozeada e /si/ com vogal vozeada. Para efeito de comparação, S2 só produziu /si/ desvozeado e S3 e S4 só produziram /si/ vozeado. Em relação à /su/, apenas um informante produziu o par desvozeado e vozeado.

É importante lembrar que nosso objetivo é investigar se a vogal átona é eliminada ou desvozeada através da medição acústica da sequencia C1-V, se houver coprodução. Se a vogal átona não é excluída, é provável que ela afete as medidas do centróide da fricativa. Tais efeitos são muitas vezes localizados em determinadas regiões espectrais. Por esta razão, as medidas do centróide foram realizadas em duas diferentes faixas de filtragem. Na primeira, as fricativas foram filtradas a 500Hz e na segunda, entre 4KHz a 9KHz. Com essa filtragem, podemos afirmar se as diferenças se manifestam mais claramente nas frequências menores e maiores. As comparações entre os centróides de todo o ruído que antecede a vogal vozeada e a vogal desvozeada para /si/ podem ser vistas na Figura 18, para filtro de 500Hz, e na Figura 19, para filtro de 4KHz a 9KHz.

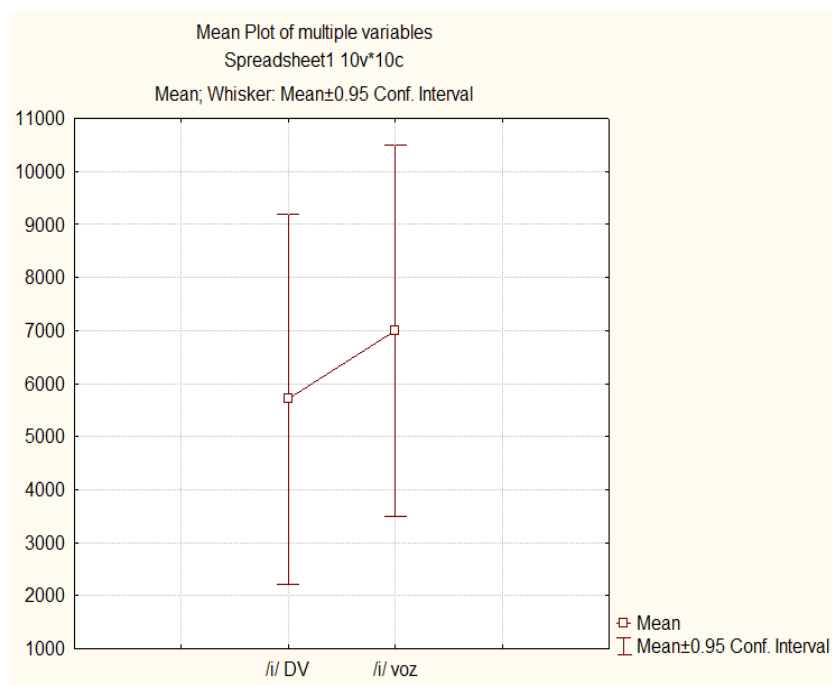


Figura 18: Centróide (Hz) /s/ que acompanha a vogal /i/ desvozeada e /i/ vozeada com filtro de 500 Hz.

Tabela 6: Teste T pareado para valores do centróide de /i/ desvozeado e /i/ vozeado.

	Média (Hz)	Grau de Liberdade	(t)	P
/i/ DV	5709,90	2	-3.8288	0.0619
/i/ Voz	6993,7			

A tabela acima mostra que o centróide não diferencia o ruído entre vogal vozeada e vogal desvozeada para uma filtragem de 500Hz. O valor de t calculado é igual a -3.8288, com $p < 0.0619$, ou seja, a diferença observada é apenas marginalmente significativa.

O filtro de 500Hz não se mostrou útil no estudo de nossos dados, precisamente porque os efeitos de diferentes vogais no espectro da fricativa parecem mais evidentes em frequências mais altas, como mostra a Figura 18. Na verdade, o [s]

afeta diretamente as baixas frequências, evitando, assim, diferenças relevantes no filtro de 500Hz.

Para um filtro de 4 kHz a 9 kHz, pode-se ver a diferença entre /i/ vozeado contra /i/ desvozeado, tendo esta última um valor médio de centróide menor.

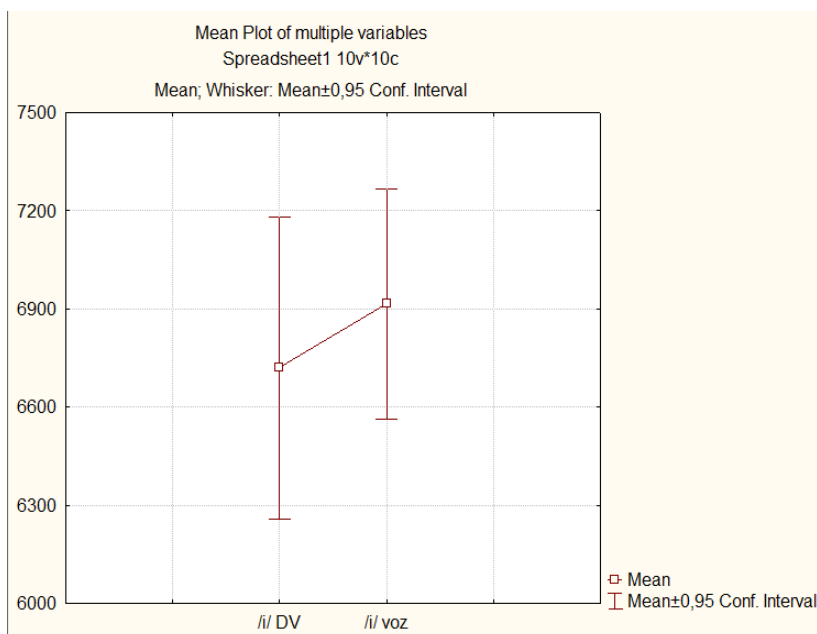


Figura 19: Centróide (Hz) do /s/ que acompanha a vogal /i/ desvozeada e /i/ vozeado.

Tabela 7: Teste T pareado para valores do centróide de /i/ desvozeado e /i/ vozeado.

	Média (Hz)	Grau de Liberdade	(t)	P
/i/ DV	6719,4	2	-6.6409	0.0219
/i/ Voz	6915			

O gráfico acima mostra que o centróide parece diferenciar o ruído que antecede a vogal vozeada (sem sinal de vogal) e o ruído característicos da vogal desvozeada (com sinal de vogal). O valor de t calculado é igual a -6.6409, com $p <$

0.0219, ou seja, a diferença observada é estatisticamente significativa. Como o valor de t é negativo, deduz-se que a média de /i/ desvozeado é inferior àquela obtida em /i/ vozeado.

A análise do Teste T pareado mostrou que o centróide parece ser um bom parâmetro para se avaliar pistas de vogal no ruído, uma vez que os valores de centróide da vogal desvozeada e vozeada apresentam-se diferentes entre si, como se observa no gráfico da Figura 19.

Assim como em Tsuchida (1994), o menor valor do centróide nossas análises pareceu evidenciar a existência de alguma pista de vogal no ruído. Alguma propriedade presente no ruído do /si/ desvozeado dos três informantes parece se diferenciar do ruído de /s/ acompanhado da vogal vozeada. Presume-se, portanto, que esta propriedade seja a presença da vogal encoberta, consequentemente surda.

Mais uma vez, a diferença que relatamos aqui advoga a favor de uma interpretação não categórica em torno do desvozeamento vocálico. Isto é, advoga a favor da análise de ensurdecimento em detrimento à queda da vogal, uma vez que as diferenças entre o centróide do ruído das fricativas não deveriam existir se as pistas das vogais também não existissem.

5.3 VSA e FCR e a redução vocálica

O enfraquecimento/redução de segmentos é caracterizado por afrouxamento articulatório, isto é, uma redução da gama de movimentos articulatórios, na medida em que o local pretendido e grau de constrição do trato vocal não são totalmente alcançados (KENT, KIM, 2003). Este “afrouxamento” resulta na centralização dos formantes da vogal, ou seja, formantes que normalmente têm frequências altas tendem a ter baixas frequências, e formantes que normalmente têm frequências baixas tendem a ter frequências mais altas.

A redução vocálica está diretamente relacionada ao ritmo e ao acento. No inglês, por exemplo, processos de enfraquecimento da sílaba, como a redução ou eliminação vocálica em sílabas átonas são comuns. A tendência a reduzir vogais em sílabas átonas é uma das características mais importantes nesta língua. Segundo Lindblom (1963), a redução vocálica é considerada característica de línguas acentuais, como o próprio inglês e o sueco. Para Lindblom (1963), há uma tendência de as vogais se aproximarem da vogal central em contextos de sílabas átonas. Stetson (1951) já observava que no inglês era possível notar uma série regular de vogais reduzidas que terminam no schwa. Vogais são mais centralizadas em seu espaço em diferentes contextos consonantais (STEVENS, HOUSE, 1963), além de serem afetadas por outras vogais (ÖHMAN, 1966; FOWLER, 1981).

No PB, segundo Albano (1999, p. 40), “há processos de redução de todas as vogais nas posições átonas, especialmente a pós-tônica”. O fato de a redução afetar também o /a/, vogal que não está sujeita à neutralização, revela a sua existência independente. Para a autora, a redução parece ser manifestação da tendência do PB a compactar o espaço vocálico nas posições átonas.

Em nosso estudo, tanto os dados de duração quanto as medidas espectrais já mostraram pistas vocálicas em contexto de desvozeamento. Ou seja, a vogal não sofre síncope, mas seus gestos são sobrepostos. A hipótese de sobreposição gestual, que já foi corroborada pelos dados até agora, pode ser complementada com outra, que explica a frequente redução da vogal em contextos favoráveis ao desvozeamento: a vogal encolhe e se sobrepõe quase inteiramente à consoante.

Demonstrado o caráter gradiente da redução e sustentada a hipótese da sua existência para a posição pós-tônicas, a FonGest dispõe de um mecanismo muito conveniente para explicar a redução das vogais: trata-se da redução da magnitude dos gestos articulatórios. Nas vogais reduzidas, encurta-se a diferença entre a posição inicial da variável do trato envolvida e o comprimento de repouso do oscilador através da qual se modela o sistema dinâmico. Com isso o gesto perde amplitude e não chega a alcançar o alvo (LINDBLOM, 1963).

A redução da magnitude do gesto articulatório prevê uma centralização das vogais tal como observada na Figura 20, em oposição à Figura 21:

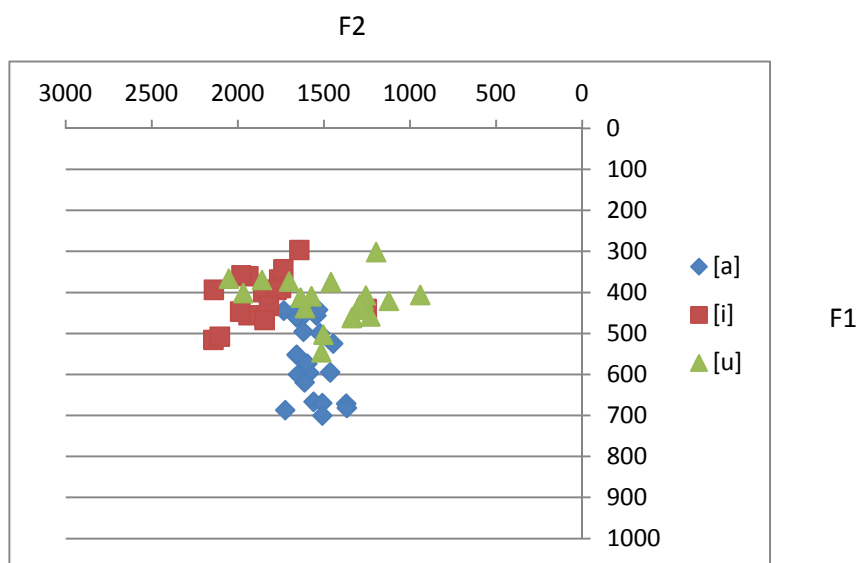


Figura 20: Espaço vocálico de vogais vozeadas em contexto de desvozeamento para sujeitos do PB.

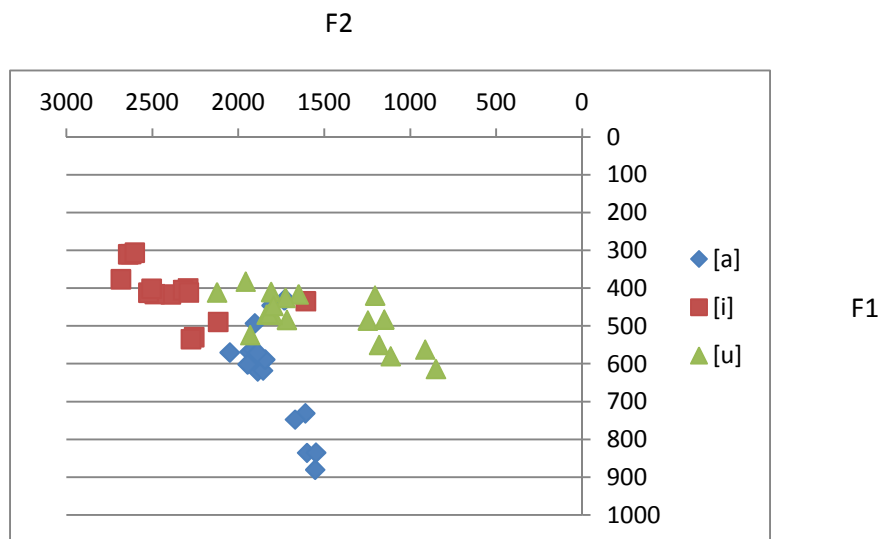


Figura 21: Espaço vocálico de vogais vozeadas em ambiente de vozeamento para sujeitos do PB.

Parece visível a compressão do espaço vocálico das vogais átonas vozeadas acompanhadas da fricativa [s] em comparação às vogais átonas em outros contextos, que, de certa maneira, já apresentam algum grau de redução.

Uma maneira usada para quantificar e representar essa centralização é com a medida da Área de Espaço Vocálico - VSA (KENT, KIM, 2003). Devido ao *undershoot* articulatório e consequente centralização vocálica, o VSA em condições de desvozeamento pode ser comprimido em relação às condições normais de produção (KENT, KIM, 2003). Assim, quando a produção de fala é clara e caracterizada pela precisão articulatória, a expansão VSA e melhora na inteligibilidade de fala (FERGUSON, KEWLEY-PORT, 2007) são esperadas. Por outro lado, certa quantidade de sobreposição articulatória e um afrouxamento na articulação produzem um VSA reduzido (cf. SAPIR *et al.*, 2010).

Além do VSA, a Razão de Centralização Formântica (FCR) é uma nova medida para a avaliação das distâncias relativas entre as vogais de diferentes falantes (SAPIR *et al.*, 2010). Enquanto razão, o FCR é projetado de modo que as frequências dos formantes do numerador são susceptíveis ao aumento, e as frequências dos formantes no denominador tendem a diminuir com a centralização da vogal. Assim, a FCR deve aumentar com a centralização e diminuir com a expansão da vogal.

Em resumo, vogais centralizadas mostram possíveis efeitos coarticulatórios, enquanto vogais periféricas indicam maior precisão de articulação. Devido à sobreposição articulatória e possível centralização vocálica, o VSA pode ser reduzido e o FCR aumentado em vogais vozeadas que estão em posições de desvozeamento em relação às vogais de condições normais. Um indicativo dessa tendência pode ser vista nas figuras abaixo:

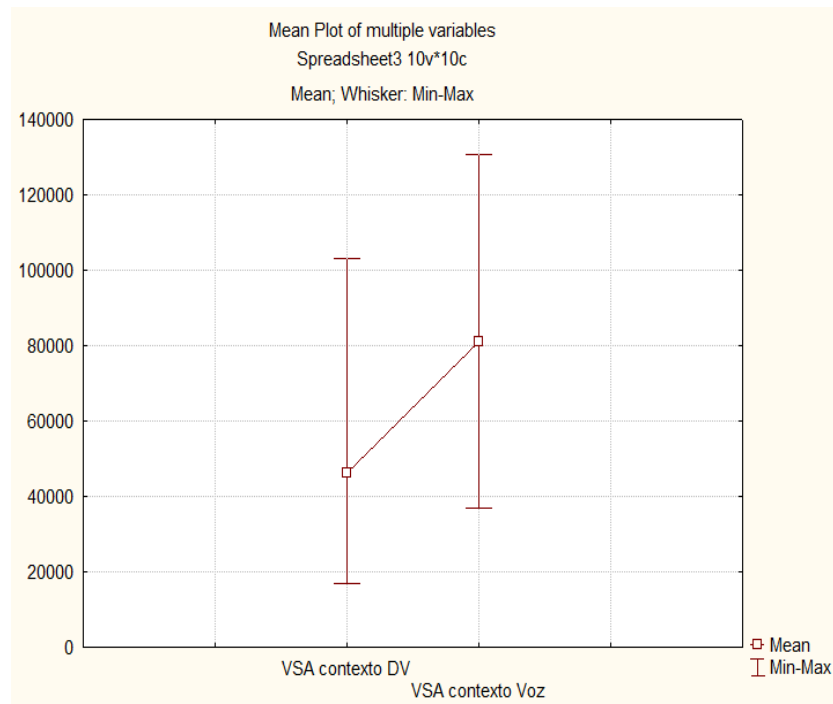


Figura 22: Valores médios da Área de espaço vocálico em contexto de desvozeamento e em contexto de vozeamento.

Como previsto, as vogais vozeadas e parcialmente desvozeadas localizadas em contextos de desvozeamento têm o VSA menor se comparadas a vogais em condições que chamamos de normal, ou seja, condições que não favorecem o desvozeamento vocálico, como mostra a figura 22 acima. No entanto, apesar dessa tendência ter sido observada nos três sujeitos deste experimento, com uma diferença relativamente grande entre as médias, as diferenças não foram significativas. Provavelmente, o número reduzido de sujeitos (apenas 3 informantes fizeram parte desta análise) impediu que a tendência observada nas médias se traduzisse em uma diferença estatística.

Para o FCR, as diferenças aparentes também corroboraram nossa hipótese inicial: as vogais sofrem uma grande redução de contextos de desvozeamento, mesmo quando vozeadas. A figura 23 abaixo mostra indícios de que a média de FCR em contexto de desvozeamento foi maior do que o mesmo em contexto de voz. Assim

como o VSA, entretanto, as diferenças entre os valores do FCR também não foram significativas.

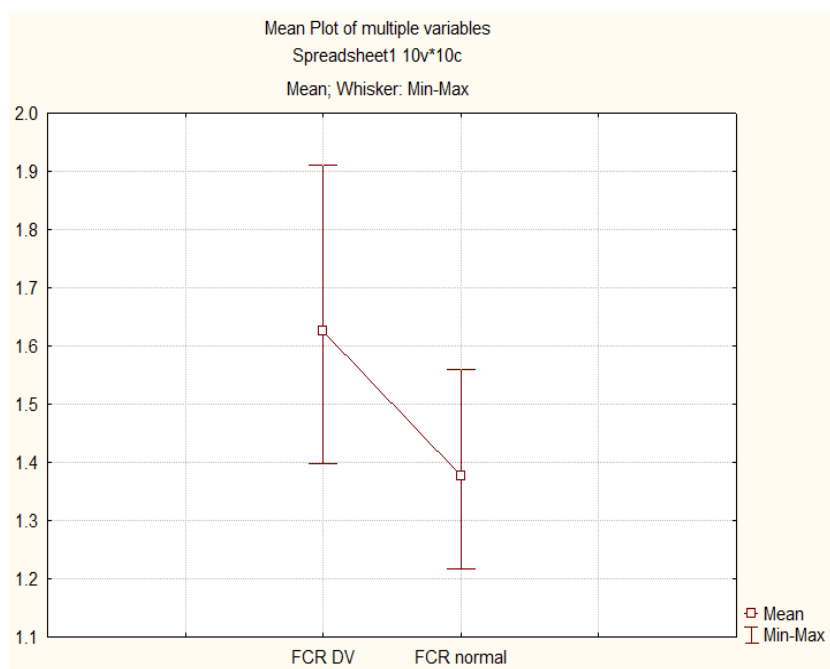


Figura 23: Valores médios da Razão de Centralização Formântica em contexto de desvozeamento e em contexto de vozeamento.

Em grande medida, podemos afirmar que a quantidade de sujeitos influenciou diretamente a não significância dos dados expostos acima. Vale salientar, o fato de todos os sujeitos apresentarem a mesma tendência: VSA menor e FCR maior para vogais em situação de desvozeamento.

Essa relação inversamente proporcional entre VSA e FCR nos parece mais clara quando observamos a Figura 24 abaixo:

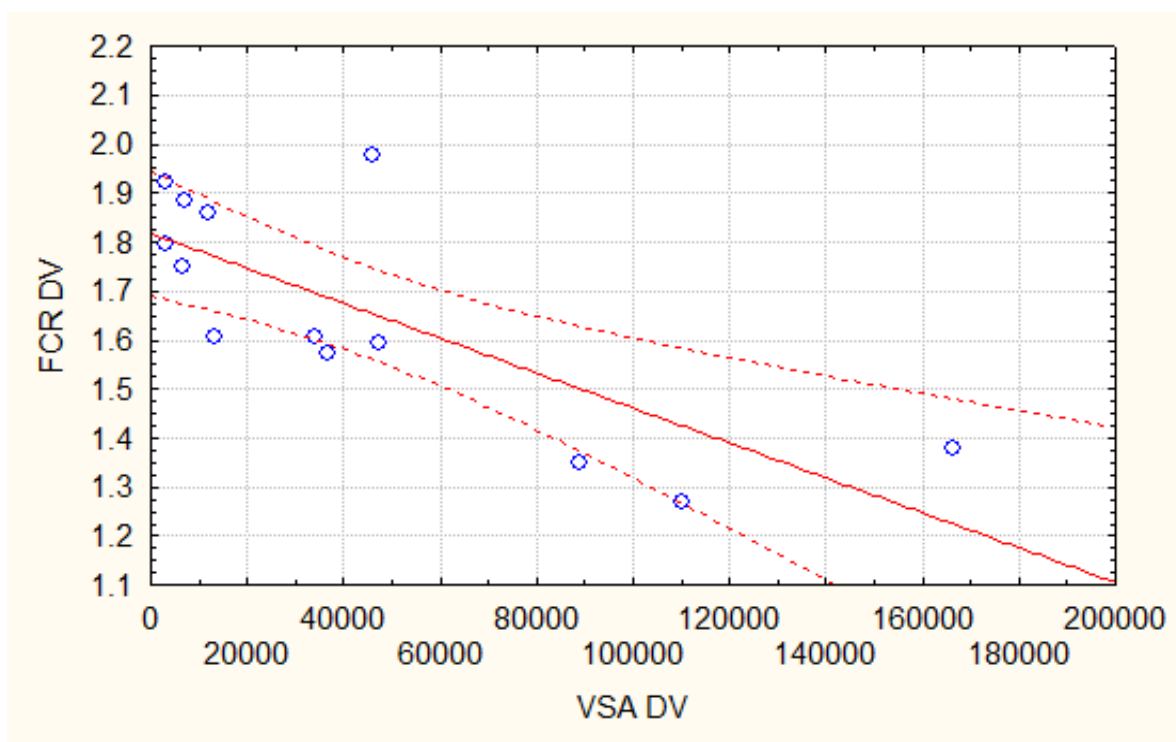


Figura 24: Diagrama de dispersão - FCR X VSA de vogais em contexto de desvozeamento.

Tabela 8: Valores de r e p para correlação entre medidas de FCR e VSA das vogais em contexto de desvozeamento

	<i>Média</i>	<i>Pares</i>	<i>r</i> (Pearson)	<i>P</i>
FCR	1.6	13	-0.77	0.002
VSA	1.3			

A correlação (-0.77) negativa alta e significativa ($p = 0.002$) existente entre o FCR e VSA só corrobora nossas impressões: essencialmente, as vogais que não são desvozeadas tornam-se mais centralizadas em condições que favorecem o desvozeamento. Essa característica pôde ser visualizada na observação direta dos espaços vocálicos, mas, só corroboradas estatisticamente com a análise da correlação dos dados de FCR e VSA.

5.4 Discussão

Inúmeros trabalhos tiveram como o foco o desvozeamento vocálico em diversos contextos. Muitos atentaram para a descrição do processo, outros ficaram restritos ao processo articulatório ou seus correlatos acústicos. Poucos, no entanto, trataram de várias características relacionadas entre si como mostramos em nossa investigação. Vale agora, apontar e comentar importantes achados de nossas análises.

Em resumo, nossos achados referem-se a quatro pontos fundamentais:

- (i) o alongamento do ruído em sílabas de vogais desvozeadas;
- (ii) o deslocamento dos centróide que diferencia sílabas com presença de vogal vozeada daquelas com vogal desvozeada;
- (iii) o desaparecimento dos formantes mais baixos da vogal sobreposta à consoantes, evidenciada pelas altos valores das medidas do centróide;
- (iv) a tendência à redução das vogais (redução da magnitude) em contexto favorecedor de desvozeamento vocálico.

Em relação à duração (i), fica claro que se a vogal realmente sofresse um processo de síncope, o esperado seria que o /s/ não apresentasse diferenças de duração nas três condições analisadas. Parece, na verdade, que ensurdecimento é resultado de uma redução temporal da parte oral da sílaba, em vez de a sílaba como um todo. O aumento observado na coarticulação em ocorrências de desvozeamento, como os dados de duração mostram, indica que a sobreposição articulatória de C sobre V pode ser a principal causa deste encurtamento da vogal e consequente alongamento da consoante, resultando no desvozeamento vocálico.

O alongamento da fricativa, conforme a FonGest, é resultado do arrastamento do período de C por V em uma relação de 1:1. Um modelo 1:1, na verdade, diz

respeito ao modo de travamento de frequência⁸ mais estável em um modelo de osciladores acoplados⁹. Um gesto de ponta da língua está associado com o ciclo do gesto de lábio, e as duas oscilações estão em uma relação de fase. O travamento de frequência de 1:1 é conhecido por ser o mais estável de um conjunto n possível (GOLDSTEIN *et al*, 2007). Em certas condições (por exemplo, as condições facilitadoras de desvozeamento), esses modos de frequência exibem transições mais simples, mais estáveis, conseqüentemente, com maior facilidade articulatória, resultando em processos como o desvozeamento.

O “afrouxamento” articulatório também está relacionado a outro resultado importante de nosso trabalho: a redução da magnitude das vogais vozeadas em ambiente de desvozeamento (iv). A pergunta que nos inquietava era qual o possível efeito dos contextos de desvozeamento sobre as vogais que eram vozeadas ou parcialmente desvozeadas? Tanto as avaliações dos gráficos de espaço vocálico, quanto as medidas de FCR e VSA encaminharam a resposta à questão: o que diferencia a produção das vogais na posição átona antecedidas por /s/ é o grau de redução da magnitude e o grau de sobreposição dessas vogais. A variação da redução e da sobreposição provoca a produção de vogais desvozeadas, parcialmente desvozeadas e vozeadas reduzidas.

Aliados aos quantificadores do espaço vocálico, a medidas de duração também alimentam a hipótese de que a redução acompanha as vogais no contexto analisado. A menor duração das sílabas parcialmente desvozeadas também é correlato da redução vocálica imposta pelas condições de desvozeamento.

É importante ressaltar que não somos os primeiros a tratar da redução em discussões em torno do desvozeamento. De acordo com Kondo (1997), há três tipos de enfraquecimentos vocálicos: (1) redução de vogal, que é a centralização da qualidade vogal com uma redução na duração, (2) ensurdecimento se vogal e, finalmente, (3) exclusão vogal, o que implica o seu desaparecimento.

⁸ O termo “travamento de frequência” refere-se ao termo “frequency locking” sem tradução específica do inglês. Em alguns textos, pode ser encontrado como “frequência de fechamento”.

⁹ O modelo de osciladores diz respeito ao segundo momento de evolução da teoria: do modelo massa-mola ao modelo de osciladores acoplados, que é o modelo atual da FonGest.

Não nos parece plausível falar na em queda na vogal nas condições analisadas e os resultados mostraram isso, uma vez que a queda implicaria inexistência do gesto vocálico. O desvozeamento, assim, seria o último estágio de uma escala de redução da magnitude vocálica, como mostramos na figura abaixo.

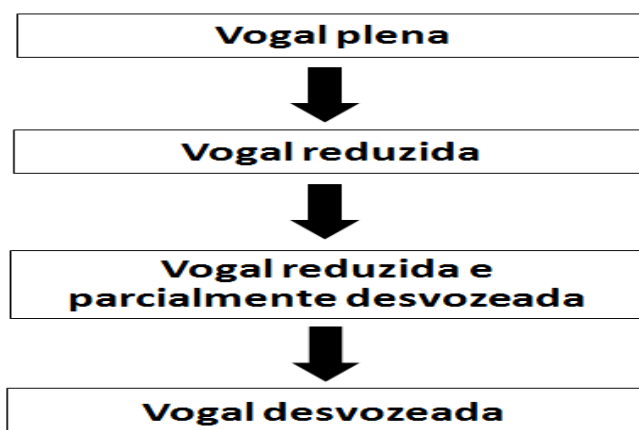


Figura 25: Proposta de redução vocálica em contexto de desvozeamento vocálico.

Entre a produção da vogal plenamente realizada, passando pela vogal reduzida até a vogal desvozeada, um declínio da duração e a redução do espaço vocálico revelam o mecanismo escalar de redução da magnitude dos gestos articulatórios.

Na FonGest, esse processo pode ser implementado por um pequeno aumento do grau de constrição causada pela sobreposição entre fricativa e vogal. É importante ressaltar que grande parte da fonologia do PB é gradiente e o único modelo fonológico apto a lidar com isso é a FonGest. Ao invés de postular regras que alterem a identidade do segmento analisado, ela altera apenas as relações entre eles: os gestos podem reduzir a sua magnitude e/ou aumentar a sua sobreposição, de tal forma que os seus resultados acústicos desapareçam ou soem alterados, como ocorre em nossos dados.

Conforme Albano (1999, pág.40), sobre a redução de vogais no PB,

intuitivamente, a neutralização e a redução parecem ser manifestações da mesma tendência do PB a compactar o espaço

vocálico nas posições átonas. Essa tendência é, entretanto, impossível de captar unificadamente nos modelos fonológicos tradicionais, que, ao conceber os traços como binários, privativos ou mesmo escalares, limitam severamente o número de distinções vocálicas possíveis, inviabilizando a expressão dos múltiplos graus de compressão do espaço vocálico envolvidos na redução.

A presença ou não de vozeamento na produção da vogal depende da maior ou menor sobreposição entre o gesto consonantal e o gesto vocálico, podendo variar de acordo com o contexto prosódico, segmental ou mesmo pragmático. Além disso, nossos dados sugerem que o desvozeamento não se difunde de maneira categórica e sim gradiente.

Em relação à medida espectral (ii), pudemos observar que o centróide (1º momento espectral) do /si/ desvozeado é diferente do ruído de /s/ acompanhado de /i/ vozeado. Essa diferença, presumimos, deve-se à presença da vogal encoberta, uma vez que as diferenças entre o centróide do ruído das fricativas não deveriam existir se as pistas das vogais tivessem desaparecido. Essa diferença corrobora os resultados de Tsuchida (1994), nos quais a autora afirma que centros de gravidade mais baixos fornecem pistas de presença de vogal. Além de Tsuchida (1994), para Torreira e Ernertus (2010), o aumento da coarticulação de /s + V/ na sílaba desvozeada resulta em um deslocamento para baixo dos parâmetros espectrais se comparado a /s/ em sílaba de vogal vozeada.

Ainda sobre os efeitos sobre o centróide, a nossa escolha do [s] como consoante que antecedia as vogais nos mostrou uma gama de resultados acústico-articulatórios. Sabe-se que uma fricativa alveolar é produzida com um estreitamento de uma constrição supraglótica que é formado por um dos articuladores da cavidade oral. Mais especificamente, é uma das classes de som produzida através do levantamento da lâmina língua contra o céu da boca, formando uma constrição e um ruído característicos das fricativas. Na produção das vogais desvozeadas, o trato

vocal é radicalmente constricto¹⁰ e o resultado acústico é a perda de energia e consequentemente o desaparecimento da estrutura formântica da vogal nas frequências mais baixas, como mostra a figura 26.

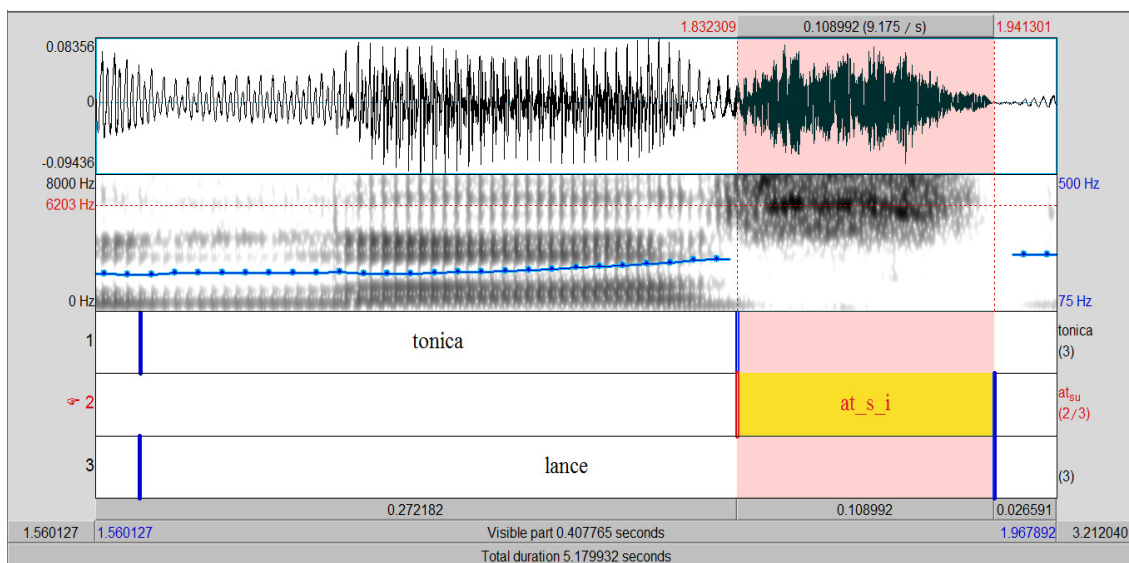


Figura 26: Espectrograma de [si] desvozeado com banda de frequência de 0-8 KHz.¹¹

Sabe-se que para o [s], o limite de frequência inferior está na casa do F4 da vogal /i/ em condições normais (VAISSIÈRE, 2009; KENT, READ, 2002).¹² Em decorrência de um maior grau de constrição no trato vocal, esse limite pode aumentar (VAISSIÈRE, 2009). Para a fricativa em [si], por exemplo, esse limite inferior está na casa dos 6KHz. Abaixo dessa linha de frequência a energia praticamente inexistente.

¹⁰ A constrição radical do trato vocal gera o que se chama de zeros ou antiformantes. Basicamente, zeros são formados sob duas condições: (i) o trato vocal está bifurcado (como em sons nasais); (ii) o trato vocal está radicalmente constricto em algum ponto. Por esta última razão, a produção de vogais desvozeadas que resultam na perda de energia na transmissão (FANT, 1960; KENT E READ, 2002).

¹¹ O espectrograma da Figura (26) tem uma janela de frequência de 0 a 8KHz. Essa janela permite a visualização da energia nas frequências mais altas e a filtragem exercida pelo [s] nas baixas frequências. Vale lembrar que as demais figuras de nosso trabalho têm janelas de frequência até 5KHz.

¹² O limite de frequência mais baixa do [j], por exemplo, está região do F3 vocálico.

Nesse sentido, justificam-se também os valores do centróide da sílaba [si] em nossos resultados. Por conta da subtração as baixas frequências, o [s] tende a jogar os valores mais baixos do centróide para cima e o mais altos para baixo. O resultado, então, foi valores do centróide do ruído de [s] nas sequencias desvozeadas de aproximadamente 6KHz.

Tudo que afirmamos até aqui, aliado à reflexão em torno da teoria fonológica que tomamos como base, fez-nos refletir sobre algo que aparece quase como consenso na literatura: a sobreposição do gesto glótico como condição para o desvozeamento vocálico.

É importante salientar que a configuração da glote em um instante no tempo não nos permite especificar se a produção é vozeada ou desvozeada. Inicialmente, a fonação depende da interação de uma série de fatores: os ajustes musculares das pregas vocais que determinam se e em que grau elas são abduzidas ou aduzidas; a pressão subglótica, que é sensível a fatores respiratórios; a elasticidade intrínseca das pregas vocais. Enfim, durante a fonação, a variação em qualquer destes fatores irá também afetar o seu modo de vibração.

Somam-se às variações articulatórias, as particularidades da produção da fricativa surda. Segundo Löfqvist e McGarr (1987), o gesto de uma fricativa surda é maior devido a aerodinâmica da produção da fricativa, em que uma grande abertura glótica não só impossibilita o vozeamento, mas também reduz a resistência ao fluxo de ar da laringe e auxilia na acumulação de pressão necessária para a produção da fonte de ruído.

Além disso, conforme Saltzman e Munhall (1989), as fricativas e, em especial o [s], dominam o comportamento da glote mais fortemente do que outras consoantes¹³. Sendo assim, não há sobreposição gestual na glote, mas sim uma dominância do gesto na consoante sobre o gesto vocálico, que não é sobreposto, pois não é realizado, uma vez que não há fechamento da glote. (Ver Figura 2 sobre resultados de Yoshioka, 1981).

¹³ A ideia de dominância do gesto fricativo sobre a glote é desenvolvida também em Goldstein (1990).

Em resumo, o gesto glótico da vogal deixa de existir, mas os gestos orais da vogal ainda estão presentes, mas são sobrepostos. As pistas obtidas em nossa investigação, a saber, a maior duração da fricativa em sílabas de vogais desvozeadas, a queda do centróide que diferencia sílabas com presença de vogal vozeada daquelas com vogal desvozeada e a tendência à redução da magnitude vocálica permitiram que assumíssemos essa conclusão.

Considerações Finais

"Ciência é conhecimento organizado. Sabedoria é vida organizada."

Immanuel Kant

O nosso trabalho atentou para fenômenos que requerem atenção e análises minuciosas. O estudo mais detalhado de alofonias gradientes, como a que investigamos, contribui para melhores argumentos em favor de quadros teóricos, como a FonGest (BROWMAN, GOLDSTEIN, 1992 e seguintes; GOLDSTEIN, FOWLER, 2003). A adoção dessa plataforma teórica, que tem se mostrado eficiente explicando processos gradientes e variáveis, trouxe esclarecimentos importantes para a explicação de nossos dados. Podemos, agora, apontar os ganhos trazidos pela perspectiva dinâmica adotada neste estudo.

Em relação à natureza do processo, os resultados reportados em outras línguas se repetiram nos dados de nossa investigação. As vogais altas são produzidas sem a vibração das pregas vocais em ambientes consonantais surdos. Como esperado, a ocorrência de vogais baixas desvozeadas foi pequena e variou muito entre os sujeitos. Como as vogais altas são mais breves do que as vogais

baixas em todas as línguas, há uma maior probabilidade de que as abduções glóticas de consoantes surdas adjacentes impeçam a plena realização das aduções glóticas necessárias para a sonoridade vocálica. Além disso, a constrição oral estreita associada à produção de vogais altas aumenta a pressão de ar na cavidade oral que inibe o fluxo de ar transglótico e, por conseguinte, torna as vogais mais susceptíveis ao ensurdecimento do que aquelas articuladas com posições mais baixas da língua.

A ocorrência de vogais totalmente e parcialmente desvozeadas e vozeadas também variou entre os informantes. Essa variabilidade, conforme a FonGest, parece ser fruto de diferenças na coordenação dos gestos vocálico e consonantal entre os informantes. A ocorrência do desvozeamento, completo e parcial, em contextos semelhantes e para um mesmo sujeito, aparentemente, deve ser condicionada pelos mesmos fatores e o limite entre a ocorrência de um ou do outro parece mínimo.

Aparentemente, um maior ou menor grau de sobreposição pode implicar na vibração mínima ou na não vibração das pregas vocais. A ocorrência de desvozeamento total e parcial e de vozeamento nos dados analisados também fortaleceu nossa ideia de que o desvozeamento não pode ser entendido como um processo categórico.

Assim como grande parte da fonologia do PB, o desvozeamento de vogais é gradiente e o único modelo fonológico apto a lidar com isso é a FonGest. O fato de as ocorrências variarem minimamente mostra que a mudança não é categórica [+ voz] para [- voz], mas um processo gradual. Em relação à representação tradicional, que lidam com operações pautadas na distintividade entre os segmentos, a Fonologia Gestual é capaz de expressar gradientes mínimos como na produção das vogais desvozeadas.

Em relação à análise fonética, que tinha por objetivo “rastrear” o processo articulatório envolvido na produção de vogais surdas, percebemos que há na produção de sílabas de vogais desvozeadas um “*algo mais*” que impossibilita a afirmação de que o sinal vocálico deixou de existir. Tanto as medidas de duração

quanto as medidas do centróide mostraram que pistas remanescentes do gesto vocálico permanecem na porção aperiódica do ruído das sílabas desvozeadas.

Na medida espectral, pudemos observar que o centroide é deslocado para baixo quando o ruído fricativo acompanha /i/ desvozeado. Essa diferença, ao que parece, deve-se à presença da vogal, mesmo encoberta. Os dados de duração, por sua vez, mostraram que se a vogal realmente sofresse um processo de síncope, o esperado seria que o /s/ não apresentasse diferenças de duração nas três condições analisadas, como efetivamente mostrou.

Os dados de FCR e VSA acrescentaram algo de novo e de diferenciado à nossa análise: a diferença entre a produção das vogais (desvozeada ou não) na posição átona é o grau de redução da magnitude da produção vocálica. Grande responsável pelo desvozeamento das vogais altas, o acento está diretamente ligado à redução e, conseqüentemente, ao desvozeamento.

Em conclusão, tanto a duração quanto o centróide e o FCR/VSA mostraram que, na nossa interpretação, a vogal desvozeada é o último estágio da redução vocálica na posição tônica do PB. Na produção da vogal vozeada, passando pela vogal parcialmente desvozeada até a vogal desvozeada, parece haver um declínio da duração, uma redução do centróide e a redução do espaço vocálico que revelam o mecanismo escalar de redução da magnitude dos gestos articulatórios. Todo esse mecanismo é gradiente e variável como mostramos em nosso estudo.

Sabemos que o estudo de fenômenos de variação em curso requer a análise de muitos falantes. A amostra, no entanto, não impediu que resultados importantes fossem obtidos por nossa investigação. Vale ressaltar que maiores estudos são necessários para sedimentar os achados desta dissertação, especialmente elucidar questões que não abordamos como: o que acontece com as vogais altas nas demais posições não acentuadas? O desvozeamento tem a mesma natureza se a vogal estiver acompanhada de oclusivas surdas? E a mudança da estrutura silábica (CV para CVC) pode mudar a ocorrência do fenômeno? Deixamos, portanto, o desejo de

continuar este trabalho, visando completar as possíveis lacunas e ampliar o estudo em torno do desvozeamento vocálico.

Referências Bibliográficas

- ALBANO, E. C. O Português Brasileiro e as controvérsias da Fonética atual: pelo aperfeiçoamento da Fonologia Articulatória. In: **D.E.L.T.A.**, São Paulo, v. 15, p. 23-50; 1999.
- _____. **O Gesto e suas Bordas**: esboço de Fonologia Acustico-Articulatória do Português Brasileiro. Campinas: Mercado de Letras/ São Paulo: FAPESP; 2001.
- ANDREEVA, B.; KOREMAN, J. The status of vowel devoicing in Bulgarian: phonetic or phonological? In: ZYBATOW, T. et al. **Formal Description of Slavic Languages**: The Fifth Conference (FDSL5). Frankfurt Main: Peter Lang; 2008.
- BATTISTI, Elisa ; HERMANS, Ben . Fixed and variable properties of Brazilian Portuguese palatalization. In: **Phonetics and Phonology in Iberia 2007**. Braga, Portugal. PaPI'07. Braga : Ed.UMinho, p. 21-22; 2007.
- BECKMAN, M.; SHOJI, A. Spectral and perceptual evidence for CV coarticulation in devoiced /si/ and /syu/ in Japanese. **Phonetica**, 41:61–71; 1984.
- BECKMAN, M. Segment duration and the 'mora' in Japanese. **Phonetica** 39, 113-135; 1982.
- _____. When is syllable not a syllable? In: OTAKE, T.; CUTLER, A. **Phonological Structure and Language Processing**: Cross-Linguistic Studies, pp. 95-123. Nova York; 1996.

- BERTI, L. C. **Aquisição incompleta do contraste entre /s/ e /ʃ/ em crianças falantes do português brasileiro**. Tese de Doutorado em Linguística. Instituto de Estudos da linguagem, Universidade Estadual de Campinas, Campinas; 2006.
- BLOCH, B. Studies in colloquial Japanese IV: Phonemics. **Language** 26, 86-125; 1950.
- BOERSMA, P.; WEENINK, D. **Praat software**. Versão 5.12. The Netherlands, Amsterdam; 2002.
- BROWMAN, C. P.; GOLDSTEIN, L. Towards an articulatory phonology. **Phonology Yearbook**, 3, p. 219–252. 1986.
- BROWMAN, C. P., GOLDSTEIN, L. Articulatory gestures as phonological units. **Phonology**, 6, p. 201–251. 1989.
- BROWMAN, C. P.; GOLDSTEIN, L. Tiers in articulatory phonology, with some implications for casual speech. In: KINGSTON, J.; BECKMAN, M. E.: **Papers in laboratory phonology: Between the grammar and the physics of**. Cambridge: Cambridge University Press. 1990
- BROWMAN C.P. & GOLDSTEIN L. **Articulatory phonology: an overview**. Haskins Laboratories, New Haven, 1992.
- BROWMAN, C. P., GOLDSTEIN, L. The Gestural Phonology Model. In: HULSTIJN, W.; PETERS, H. F. M.; VAN LIESHOUT, P.H.H. M. **Speech production: Motor control, brain research and fluency disorders**. Amsterdam, Holanda: Elsevier. 1997.
- CEDERGREN, H. La chute des voyelles hautes en français de Montreal: Astu entendu la belle syncope?. In: LEMIEUX, M.; CEDERGREN, H. **Les tendances dynamiques du français parlé à Montreal**, pp. 57–145. 1985.
- CHITORAN, I.; ISKAROVS, K. Acoustic evidence for high vowel devoicing in Lezgi. In: SOCKS et al. **Proceedings of ISSP 8**, Strasbourg, France. 93-96. 2008.
- CHISTOVICH, L.A.; SHEIKIN, R.L. LUBLINSKAJA, V.V. “Centers of Gravity” and Spectral peaks at the determinants of Vowel Quality. In: LINDBLOM, B.; OHMAN, S. In: **Frontiers of Speech Communication Research**. pp.143-157. 1979.

- CHISTOVICH, L.A.; LUBLINSKAJA, V.V. The “Center of Gravity” effect in vowel spectra and critical distance between the formants: Psychoacoustics study of perception of vowel-like stimuli. **Hearing Research** 1. 185-195. 1979.
- CRUZ-FERREIRA, M. Portuguese (European). In: **Handbook of the International Phonetic Association**. Cambridge: Cambridge University Press, pp. 126–129.1999.
- DAUER, R. M. The reduction of unstressed high vowels in modern Greek. **Journal of the International Phonetic Association**, 10:17–27. 1980.
- DAVIDSON, L. Schwa Elision in Fast Speech: Segmental Deletion or Gestural Overlap? **Phonetica**. pp. 63:79. 2006.
- D’ECHAINE, R.M. Stress in Québécois: evidence from high vowels. In: DOBRIN, L.; NICHOLS, L.; RODRIGUEZ, R. **CLS 27: Part I: the General Session**. Chicago Linguistic Society, pp. 107–118. 1991.
- DELFORGE, A. Gestural Alignment Constraints and Unstressed Vowel Devoicing in Andean Spanish. **Proceedings of the 26th West Coast Conference on Formal Linguistics**, Somerville, MA: Cascadilla. 2008.
- FAGYAL, Z.; MOISSET, C. Sound change and articulatory release: where and why are high vowels devoiced in Parisian French. In: OHALA et al. **Proceedings of the XIVth International Congress of Phonetic Sciences**. Berkeley: University of California, pp. 309–312. 1999.
- FANT, G. **Acoustic theory of speech production**. Mouton: The Hague; 1960.
- FERGUSON S.; KEWLEY-PORT, D. Talker differences in clear and conversational speech: acoustic characteristics of vowels. **Journal of Speech-Language and Hearing Research** 50: 2007.
- FORREST, K.; WEISMER, G.; MILENKOVIC, P. DOUGALL, R.N. Statistical Analysis of word-initial voiceless obstruent: preliminary data. **JASA** 84, 115-123. 1988.
- FOWLER, C. Production and perception of coarticulation among stressed and unstressed vowels. **J. Speech Hear. Res.** 46:127-139. 1981.
- FUJIMOTO, Masako. Vowel Duration and Vowel Devoicing in Japanese : A Comparison between Tokyo and Osaka Dialect Speakers. The Society of Japanese Linguistics. **Kokugogaku : studies in the Japanese language** 55(1) pp.2-15. 2004

- GOLDSTEIN, L. On articulatory binding: Comment's on Kingston's paper. In J. KINGSTON, J.; BECKMAN, M.E. (eds.), **Papers in laboratory phonology**: 1. Between the grammar and physics of speech, pp. 445-450. Cambridge, CUP. 1990.
- GOLDSTEIN, L.; FOWLER, C. Articulatory Phonology: a phonology for public language use. In: N. O. SCHILLER; A. MEYER (orgs.), **Phonetics and Phonology in language comprehension and production**: differences and similarities. Berlin; 2003.
- GOLDSTEIN, L.; POUPLIER, M.; CHEN, L.; SALTZMAN, E.; BYRD, D. Dynamic action units slip in speech production errors. **Cognition** 103, 386–412; 2007.
- GORDON, Matthew. The phonetics and phonology of non-modal vowels: A cross-linguistic perspective. **Proceedings of the Berkeley Linguistics Society** 24: 93-105. 1998.
- GREENBERG, J. Some methods of dynamic comparison in linguistics. In: PUHVEL, J. (Ed.), **Substance and structure of language**, pp. 147–204. Los Angeles: 1969. Center for research in languages and linguistics
- HASEGAWA, Yoko. Pitch accent and vowel devoicing in Japanese. **Proceedings of the 14th International Congress of Phonetic Sciences**, 523-26, 1999
- HAN, M. S. **Unvoicing of vowels in Japanese. Onsei no Kenkyu** (Study of Sounds), Nihon Onsei Gakkai (The Phonetic Society of Japan) 10, 81-100. 1962.
- HAYES, B. Sample Paper: High Vowel Devoicing in Québécois. **Linguistics** 120: Winter.2001.
- HELGASON, P.; KOHLER, K. J. Vowel deletion in the Kiel Corpus of Spontaneous Speech. In KOHLER, K. J.; REHOR, C.; SIMPSON, A. P. **Sound Patterns in Spontaneous Speech, AIPUK** 30. 1996.
- HIRAYAMA, Manami. **Postlexical prosodic structure and vowel devoicing in Japanese**. Ph.D. dissertation. Department of Linguistics.University of Toronto. 2009.
- HIROSE, H. The activity of the adductor laryngeal muscles in respect to vowel devoicing in japanese. **Phonetica**, 23:156–170. 1971.
- JAEGER, A. Speech aerodynamics and phonological universals. In: **Proceedings of the annual meeting of the Berkeley Linguistics Society**, Vol. 4. 1978.

- JANNEDY, S. Gestural phasing as an explanation for vowel devoicing in Turkish. **Ohio State University Working Papers in Linguistics**, 45:56–84. 1995.
- JOHNSON, K. **Acoustic & Auditory Phonetics**. 2.ed. Oxford: Blackwell, 2003.
- JUN, S-A.; BECKMAN, M. **A gestural-overlap analysis of vowel devoicing in Japanese and Korean**. Annual Meeting of the Linguistic Society of America, Los Angeles, January, 1993.
- JUN, Sun-Ah, BECKMAN, Mary, H.-J. L. Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for korean obstruents. **UCLA Working Papers in Phonetics**, 96:43–66. 1998.
- KAWAKAMI, S. **An outline of Japanese sounds**. Tokyo. 1977.
- KENT, R.; KIM, Y. Toward an acoustic typology of motor speech disorders. **Clinical Linguistics and Phonetics** 17: 427–445. 2003.
- KENT, R. D.; READ, C. **The Acoustic Analysis of Speech**. San Diego: Singular, 1992.
- KONDO, M. **Mechanisms of Vowel Devoicing in Japanese**. Ph.D. dissertation. University of Edinburgh. 1997.
- _____. Syllable structure and its acoustic effects on vowels in devoicing environments. In: VAN DE WEIJER, J.; NANJO, K.; NISHIHARA, T. **Voicing in Japanese**. Berlin. 2005.
- LAVER, J. **Principles of Phonetics**. Cambridge: Cambridge University Press. 1994.
- LEHISTE, I. **Suprasegmentals**. MIT, Cambridge: Massachusetts, 1970.
- LEMLE, Miriam. **Phonemic System of the Portuguese of Rio de Janeiro**. Dissertação de Mestrado. Universidade da Pensilvânia. 1966.
- LINDBLOM, B. Spectrographic study of vowel reduction. **J. Acoust. Soc. Am.** 35: 1963.
- LINDBLOM, B. E. e SUNDBERG, J. E., Acoustical consequences of lip, tongue, jaw, and larynx movement, **Journal of the Acoustic Society of America**, 50(4): 1166-1179, 1971.

- LÖFQVIST, A.; MCGARR, N. Laryngeal dynamics in voiceless consonant production. In BAER, T.; SASAKI, C.; HARRIS, K.S. **Laryngeal function in phonation and respiration**. Boston: College Hill. 1987.
- MAEKAWA, K. Production and perception of the accent in the consecutively devoiced syllables in Tokyo Japanese. In: **Proceedings of ICSLP 90**. pp. 61-64. 1990.
- MAEKAWA, K.; H. KIKUCHI. Corpus-based analysis of vowel devoicing in spontaneous Japanese: an interim report. In J. VAN DE WEIJER, K. Nanjo, and T. NISHIHARA (Eds.), **Voicing in Japanese**, pp. 205-228. Berlin; New York: Mouton de Gruyter. 2005.
- MCCAWLEY, J. D. The **Phonological Component of a Grammar of Japanese**. The Hague: Mouton Co. N.V., Publishers. 1968.
- MO, Y. Temporal, spectral evidence of devoiced vowels in Korean. In **Proceedings of ICPHS 2007**, Saarbrücken. 2007.
- OHALA, J. J. A mathematical model of speech aerodynamics. In: FANT, G. (Ed.), **Speech communication: Speech production and synthesis by rule**. Vol. 2. 1975.
- OHALA, J. J. A model of speech aerodynamics. **Report of the Phonology Laboratory** (Berkeley) 1:1976.
- OHALA, J. J. The listener as a source of sound change. In: _____ et al. **Papers from the Parasession on Language and Behavior**. Chicago: Chicago Linguistic Society, 1981, p. 177-213.
- OHISO, M. A phonological study of some English loan words in Japanese. **Ohio State University Working Papers in Linguistics**. 1973.
- ÖHMAN, S. Coarticulation in VCV utterances: spectrographic measurements. **J. Acoust. Soc. Am.** 39:151-168. 1966.
- OUELLET, M.; TZOUKERMANN, E.; M'ENARD, L. High vowel /iyu/ in Canadian and Continental French: an analysis for a TTS system. **Proceedings of Eurospeech 99**, vol. 5:1999.
- RINALDI, L. M. Procedimentos para a análise das vogais e obstruintes na fala infantil do português brasileiro. Dissertação (mestrado): Universidade Estadual de Campinas, Instituto de Estudos da Linguagem. Campinas, SP: 2010.

- RODGERS, J. E. J. Vowel deletion/devoicing. In: SIMPSON, A. P. ; PATZOLD, M. **Sound Patterns of Connected Speech: Description, Models, and Explanation**, AIPUK 31, 1996.
- RODGERS, J. E. J. Vowel devoicing/deletion in English and German. In: SIMPSON A. P.; KOHLER, K. J.; RETTSTADT, T. **The Kiel Corpus of Read/Spontaneous Speech** — Acoustic data base, processing tools and analysis results, AIPUK 32. 1997.
- SALTZMAN, E.; KELSO, J. A. S. Skilled actions: a task dynamic approach. In: **Psychological Review**. n. 94, p. 84-106. 1987.
- SALTZMAN, E.; MUNHALL, K. A dynamic approach to gestural patterning in speech production, **Ecological Psychology**, 1(4), 333-382; 1989.
- SAPIR, S. et al. **Formant Centralization Ratio (FCR): A proposal for a new acoustic measure of dysarthric speech. J Speech Lang Hear Res. February ; 53(1). 2010.**
- SATO, H. Accent rules and sequential voicing rules in compounds. In: **Lectures on Japanese and Japanese Education. Vol. 2.** pp. 72-106. Cambridge: Cambridge University Press. 1989.
- SHIRAISHI, H. Vowel devoicing of Ainu: How it differs and not differs from vowel devoicing of Japanese. In: HONMA, T. OKAZAKI, M. TABATA, T. TANAKA. S. **A New Century of Phonology and Phonological Theory**. Tokyo, 237-249. 2003.
- SILVA, A. H. P. **As fronteiras entre Fonética e Fonologia e a alofonia dos róticos iniciais em PB: dados de dois informantes do sul do país.** Tese de doutorado, inédita. LAFAPE/IEL/UNICAMP, 2002.
- SMITH, C. L. Vowel devoicing in contemporary French. **Journal of French Language Studies**, 13(2):177–194. 2003.
- SOLI, S. Second formants in fricatives: Acoustic consequences of fricative-vowel coarticulation. **JASA** 70, 976-984. 1981.
- STETSON, R.H. **Motor Phonetics**. North-Holland Publishing Co., Amsterdam, 1951.
- STEVENS, K.N.; HOUSE, A. S. Perturbation of vowel articulations by consonantal context: an acoustical study. **J. Speech Hear. Res. 6.** 1963
- SUGITO, M. **Studies on Japanese Accent**. Tokio: Sanseido. 1982.

- TORREIRA, F.; ERNESTUS, M. Phrase-medial vowel devoicing in spontaneous French. **INTERSPEECH 2010**. September, Chiba, Japan, pp.26-30, 2010.
- TSUCHIDA, A. Fricative-vowel coarticulation in Japanese devoiced syllables: Acoustic and perceptual evidence. **Papers of Cornell Phonetics Laboratory**, vol. 9, pp. 183-222. 1994.
- TSUCHIDA, A. **Phonetics and phonology of Japanese vowel devoicing**. PHD dissertation, Cornell University. 1997
- VANCE, T. J. **An Introduction to Japanese Phonology**. State University of New York Press. 1987.
- VAISSIÈRE, J. Articulatory modeling and the definition of acoustic-perceptual targets for reference vowels. **The Chinese Phonetics Journal**, Vol. 2:22-33. 2009.
- YOSHIOKA, H. Laryngeal adjustments in the production of the fricative consonants and devoiced vowels in Japanese. **Phonetica**, 38:236–251. 1981.
- YOSHIDA, N.; SAGISAKA, Y. **Factor analysis of vowel devoicing**. ATR technical report. 1990.
- WALKER, D.C. **The Pronunciation of Canadian French**. Ottawa: University of Ottawa Press. 1984.

Anexos

Anexo A: Lista de Palavras

<i>Palavra</i>	<i>Frase-veículo</i>	<i>Frequência de ocorrência</i>
<i>aço</i>	<i>O aço parecia não enferrujar muito.</i>	<i>703</i>
<i>baço</i>	<i>O baço parecia precisar de atenção.</i>	<i>52</i>
<i>masso</i>	<i>O masso parecia muito mais pesado.</i>	<i>94</i>
<i>lance</i>	<i>O lance parecia perfeito no momento.</i>	<i>1019</i>
<i>passe</i>	<i>O passe parecia resultar em golaço</i>	<i>1631</i>
<i>tosse</i>	<i>A tosse parecia sintoma de gripe.</i>	<i>132</i>
<i>onça</i>	<i>A onça parecia mais fraca e faminta.</i>	<i>217</i>

<i>caça</i>	<i>A caça parecia proibida na data.</i>	<i>743</i>
--------------------	--	-------------------

<i>massa</i>	<i>A massa parecia esperar a vitória.</i>	<i>290</i>
---------------------	--	-------------------

<i>taça</i>	<i>A taça parecia destinada ao jogador.</i>	<i>786</i>
--------------------	--	-------------------

Anexo B: Script de duração

form Entre com os parametros iniciais

sentence Diretorio_dos_dados

sentence Arquivo_de_saida duracao.txt

sentence Extensao_do_arquivo_sonoro .wav

sentence Extensao_do_TextGrid .TextGrid

endform

Constants (8)

outfile\$ = arquivo_de_saida\$

dataDirectory\$ = diretorio_dos_dados\$

emptyString\$ = ""

lengthExtension = length(extensao_do_arquivo_sonoro\$)

Writing header (14)

fileappend "'outfile\$" 'objName\$' 'tab\$' 'labelSeg\$' 'tab\$' 'durUA' 'newline\$'

call AddLine

Create Strings as file list... fileList 'dataDirectory\$''extensao_do_arquivo_sonoro\$'*

select Strings fileList

numberOfFiles = Get number of strings

Start big fat loop (21)

for ifile from 1 to numberOfFiles

select Strings fileList

fileName\$ = Get string... ifile

Read from file... 'dataDirectory\$'fileName\$'

lengthFileName = length(fileName\$)

objName\$ = left\$ (fileName\$, lengthFileName-lengthExtension)

Read from file... 'dataDirectory\$'objName\$'extensao_do_TextGrid\$'

select TextGrid 'objName\$'

Number of tiers is not known in advance (32)

nTiers = Get number of tiers

for aa from 1 to nTiers

aa is the tier number. Unknown in advance. (35)

select TextGrid 'objName\$'

nIntTierTemp = Get number of intervals... aa

select TextGrid 'objName\$'

plus Sound 'objName\$'

Extract intervals preserving time. Messy but necessary (40)

Extract non-empty intervals... aa 1

One or more Sound objects are selected.

Necessary to get their names in advance? (42)

nSelectedSounds = numberOfSelected("Sound")

```

for bb from 1 to nSelectedSounds
    sound'bb' = selected ("Sound", bb)
endfor

for cc from 1 to nSelectedSounds
    labelSeg$ = selected$ ("Sound")

    if (labelSeg$ == "at_s_u" or labelSeg$ == "a" or labelSeg$ == "i" or
labelSeg$ == "u" or labelSeg$ == "atona")

        select sound'cc'

        Rename... UnderAnalysis

        durUA = Get total duration

        durUA = round (durUA*1000)

        call WriteOutput

        call CleanUP

    endif #endif match label

endfor

endfor #aa

endfor #big fat loop.

select all

Remove

#####

procedure AddLine

    fileappend "'file_name$'" 'newline$'

```

endproc

procedure WriteOutput

fileappend '''outfile\$''' 'objName\$' 'tab\$' 'labelSeg\$' 'tab\$' 'durUA' 'tab\$' 'newline\$'

endproc

procedure CleanUP

cleaning up

select Sound UnderAnalysis

Remove

endproc

Anexo C: Script de medidas do centroide:

form Entre com os parametros iniciais

sentence Diretorio_dos_dados

sentence Arquivo_de_saida

sentence Extensao_do_arquivo_sonoro .wav

sentence Extensao_do_TextGrid .TextGrid

sentence Label_dos_segmentos

endform

Constants

outfile\$ = arquivo_de_saida\$

dataDirectory\$ = diretorio_dos_dados\$

powerSpectralMoments = 2

emptyString\$ = ""

spectralMomentLabel\$ = label_dos_segmentos\$

lengthExtension = length(extensao_do_arquivo_sonoro\$)

Writing header

fileappend "'outfile\$'" 'objName\$' 'tab\$' 'labelSeg\$' 'tab\$' 'cg' 'tab\$'

fileappend "'outfile\$'" 'dp' 'tab\$' 'assm' 'tab\$'

fileappend "'outfile\$'" 'curt'

call AddLine

```

Create Strings as file list... fileList 'dataDirectory$'*'extensao_do_arquivo_sonoro$'

select Strings fileList

numberOfFiles = Get number of strings

# Start big fat loop

for ifile from 1 to numberOfFiles

select Strings fileList

fileName$ = Get string... ifile

Read from file... 'dataDirectory$'fileName$

lengthFileName = length(fileName$)

objName$ = left$ (fileName$, lengthFileName-lengthExtension)

Read from file... 'dataDirectory$'objName$"extensao_do_TextGrid$'

select TextGrid 'objName$'

# Number of tiers is not known in advance

nTiers = Get number of tiers

for aa from 1 to nTiers

    # aa is the tier number. Unknown in advance.

    select TextGrid 'objName$'

    nIntTierTemp = Get number of intervals... aa

    select TextGrid 'objName$'

    plus Sound 'objName$'

    # Extract intervals preserving time. Messy but necessary

```



```

Extract non-empty intervals... aa 0

# One or more Sound objects are selected.

# Necessary to get their names in advance?

nSelectedSounds = numberOfSelected("Sound")

for bb from 1 to nSelectedSounds

    sound'bb' = selected ("Sound", bb)

endfor

for cc from 1 to nSelectedSounds

    labelSeg$ = selected$ ("Sound")

    if (labelSeg$ == spectralMomentLabel$)

        select sound'cc'

        Rename... UnderAnalysis

        select Sound UnderAnalysis

    endfor

    total = Get total duration

    #startSel = Get start of selection

    meio = total/2

    select Sound UnderAnalysis

    Edit

    editor Sound UnderAnalysis

    Show analyses... 1 1 0 0 0 5.0

    #print 'startSel'

```

print 'meio'

Select... 0 1

Extract selected sound (time from 0)

Close

endeditor

select Sound untitled

Filter (pass Hann band)... 0 4000 9000

select Sound untitled_band

Rename... windowUA

To Spectrum... yes

select Spectrum windowUA

cg = Get centre of gravity... powerSpectralMoments

dp = Get standard deviation... powerSpectralMoments

assm = Get skewness... powerSpectralMoments

curt = Get kurtosis... powerSpectralMoments

call AddLine

call WriteOutput

select Sound windowUA

plus Spectrum windowUA

```

        else

        endif

    endfor

endfor

endfor # end of big loop


select all

Remove

#####

procedure AddLine

    fileappend "'outfile$" 'newline$'

endproc


procedure WriteOutput

    fileappend "'outfile$" 'objName$' 'tab$' 'labelSeg$' 'tab$' 'cg' 'tab$'

    fileappend "'outfile$" 'dp' 'tab$' 'assm' 'tab$'

    fileappend "'outfile$" 'curt'

endproc


procedure CleanUP

# cleaning up

select Sound UnderAnalysis

Remove

endproc

```